SCHOLARLY PUBLICATIONS

A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT

@DTU

(93rd Edition)

SEPTEMBER 2020

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 **Ninety Third** 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 **September, 2020.** 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. 6G for Bridging the Digital Divide: Wireless Connectivity to Remote Areas, Abdelaali Chaoub, Marco Giordani, ***7.1.Brejesh Lall***, Vimal Bhatia, Adrian Kliks, Luciano Mendes, Khaled Rabie, Harri Saarnisaari, Amit Singhal, Nan Zhang and Sudhir Dixit, Electronics, DTU

2. A COMPREHENSIVE OUTLINE OF DIVERSE IDS/IPS WITH RESPECT TO HYPERVISORS, ***6.A. N. SESHU KUMAR[1]***, ***8.R. KUMAR YADAV[1]***, AND ***3.N. S. RAGHAVA[2]***, [1]Computer and [2]Electronics, DTU

3. A Framework for the Improvement of Frugal Design Practices, ***3.Ravindra Singh[1]***, ***3.Sumedha Seniaray[2]*** and Prateek Saxena, [1]Design and [2]Mathematics, DTU

4. A New Robust Fuzzy Clustering Approach: DBKIFCM, Anjana Gosain and ***3.Sonika Dahiya***, Computer, DTU

5. A novel intuitionistic fuzzy co-clustering algorithm for brain images, ***3.Om Prakash Verma*** and ***6.Heena Hooda***, IT, DTU

6. A Novel MS-Word Text Document Multiple Watermarking Technique for Authentication and Copyright Protection, ***8.Ayush Arora***, ***8.Drishti Agarwal*** and ***3.Jeebananda Panda***, Electronics, DTU

7. A Route Stable Energy and Mobility aware routing protocol for IoT, ***3.Rajesh Kumar Yadav*** and ***7.Nimisha Awasthi***, Computer, DTU

8. A Survey of Application of ML and Data Mining Techniques for Smart Irrigation System, ***8.Mr. Himanshu Nandanwar***, ***3.Mrs. Anamika Chauhan***, ***8.Mr.Dushyant Pahl*** and ***8.Mr.Harshit Meena***, IT, DTU

9. An Experimental Comparison of Different Object Tracking Algorithms, ***3.NS Raghava***, ***8.Kushagra Gupta***, ***8.Ishita Kedia*** and ***8.Anirudh Goyal***, Electronics, DTU

10. Circular economy under the impact of IT tools: a content-based review, ***6.Syed Wasiul Hasan Rizvi***, ***3.Saurabh Agrawal*** and ***3.Qasim Murtaza***, Mechanical, DTU

11. Classification of EEG signal using LSTMs under Audiovisual Stimuli, ***Tejas Agarwal***, ***8.Shubham Raturi***, ***8.TK Vybhav*** and ***3.Mukhtiar Singh***, Electrical, DTU

12. COVID-19 Outbreak Prediction using Quantum Neural Networks, ***8.Pranav Kairon*** and Siddhartha Bhattacharyya, Physics, DTU

13. Designing and Performance Analysis of 15KWP Grid Connection Photovoltaic System Using Pvsyst Software, ***8.Akash Chauhan***, ***8.Mayank Sharma*** and ***3.Sarita Baghel***, Physics, DTU

14. Drag Reduction by Application of Aerodynamic Devices in a Race Car, ***8.Devang S. Nath***, ***8.Prashant Chandra Pujari***, ***8.Amit Jain*** and ***3.Vikas Rastogi***, Mechanical, DTU

15. Effect of Concentration of Alumina Nanoparticles on Performance and Emissions of Biodiesel Fueled Compression Ignition Engine, ***8.Shashank Bansal***, ***8.Yash Dagar***, ***8.Yash Kumar Singh*** and ***3.Naushad A. Ansari***, Mechanical, DTU

16. Electron plasma wave excitation by a q-Gaussian laser beam and subsequent electron acceleration, ***6.Monika Yadav***, Devki Nandan Gupta and ***3.Suresh C. Sharma***, Physics, DTU

17. Electronically Tunable Multifunction Transadmittance-Mode Fractional-Order Filter, ***7.Garima Varshney***, ***3.Neeta Pandey*** and ***3.Rajeshwari Pandey***, Electronics, DTU

18. Electrospun pH responsive poly (acrylic acid-co- acrylamide) hydrogel nanofibrous mats for drug delivery, ***6.Preeti Gupta*** and ***3.Roli Purwar***, Chemistry, DTU

19. Email Spam Detection Using Machine Learning Algorithms, ***8.Nikhil Kumar***, ***8.Sanket Sonowal*** and ***8.Nishant***, CSE, DTU

20. Enhanced Solar PV Power Generation Under PSCs Using Shade Dispersion, ***3.Mayank Kumar***, Electrical, DTU

21. Experimental system design for online characterization and performance analysis of PV module under distinguish environmental conditions, Rahul Anand, Yogesh K. Chauhan, ***3.Vinod Yadav*** and Rupendra Pachauri, Electrical, DTU

22. Formal vs informal demand for credit by women in urban India: a comparative study in slums of Delhi and Mumbai, ***6.Shruti Malik***, ***3.Girish Chandra Maheshwari*** and ***3.Archana Singh***, DSM, DTU

23. Hydrogeochemical Assessment of Groundwater for Drinking and Agricultural Use: A Case Study of Rural Areas of Alwar, Rajasthan, ***8.1Ankur Aggarwal***, ***8.1.Jigyasa Soni***, ***8.1.Khyati Sharma***, ***8.1.Mohnish Sapra***, ***8.1.Chitrakshi***, Oznur Karaca and ***3.A. K. Haritash***, Environmental, DTU

24. Impact of Graded Back-Barrier on Linearity of Recessed Gate InAlN/GaN HEMT, ***6.Megha Sharma*** and ***3.Rishu Chaujar***, Physics, DTU

25. Impact of Metal Strip on Nanoscale Double Gate Overlap Tunnel FET, ***8.Anushka Jain***, ***8.Saransh Sharma*** and ***3.Rishu Chaujar***, Physics, DTU

26. Improving Software Maintainability Predictions using Data Oversampling and Hybridized Techniques, ***3.Ruchika Malhotra*** and ***3.Kusum Lata***, CSE, DTU

27. Inpainting of Irregular Holes in a Manuscript using UNet and Partial Convolution, ***8.Amreen Kaur***, ***8.Ankit Raj***, ***3.N. Jayanthi*** and ***3.S.Indu***, Electronics, DTU

28. Introduction to Microwave Remote Sensing, ***8.Sahaj Rastogi***, Physics, DTU

29. Investigating the relation between chaos and the three body problem, ***8.T.S.Sachin Venkatesh*** and Vishak Vikranth, Physics, DTU

30. Linear Feedback Shift Register-Based Test Pattern Generators – A Comparative Study, ***8.Parangat Mittal*** and ***8.Daksh Shah***, DTU

31. LOGIC SIGNED PETRI NET, ***3.PAYAL*** AND ***3.SANGITA KANSAL***, Applied Mathematics,, DTU

32. Low Power Quaternary Adder Using CNFET, ***3.Anurag Chauhan***, ***7.1.Lokesh Mahor*** and ***3.Piyush Tiwari***, Electronics, DTU

33. MixBoost: Synthetic Oversampling with Boosted Mixup for Handling Extreme Imbalance, ***6.Anubha Kabra***, Ayush Chopraz, Nikaash Puriy, Pinkesh Badjatiyay, Sukriti Verma, Piyush Guptay and Balaji K, DTU

34. Modeling of Waste water Collection Network Using Arc GIS and SewerGEMS in Kabul, Afghanistan, ***6.Ali Reza Noori*** and ***3.S.K.Singh***, DTU

35. Numerical simulations on the aerodynamics of the Ahmed body at different slant angles, ***6.Gaurav Kumar***, Chanfiou Ahmed Mboreha and Yahya Al-fakih, Mechanical, DTU

36. ON GEOMETRICAL PROPERTIES OF CERTAIN ANALYTIC FUNCTIONS, ***3.S.SIVAPRASAD KUMAR*** AND ***6.KAMALJEET GANGANIA***, Applied Mathematics, DTU

37. Performance Analysis of a Novel Heteromaterial InAs/GaAs Junctionless TFET, ***6.Samriti Sharma*** and ***3.Rishu Chaujar***, Physics, DTU

38. PERFORMANCE MEASUREMENT OF TQM USING INTEGRATED FUZZY-AHP, ***3.RAKESH KUMAR*** and ***3.R.S. MISHRA***, Mechanical, DTU

39. Prediction and forecast for COVID-19 Outbreak in India based on Enhanced Epidemiological Models, ***8.Siddharth Singh***, ***8.Piyush Raj***, ***8.Raman Kumar*** and ***3.Rishu Chaujar***, Physics, DTU

40. QNBP NN-based I cos / algorithm for PV systems integrated with LV/ MV grid, ***6.Priyanka Chaudhary*** and ***3.M. Rizwan***, Electrical, DTU

41. Denoising SMS Text Using Ternary Tree for FAQ Retrieval, ***3.Manoj Kumar*** and ***7.Vipra Bhatt***, CSE, DTU

42. Rapid detection of biomolecules in a dielectric modulated GaN MOSHEMT, ***6.Shaveta***, ***6.H. M. Maali Ahmed*** and ***3.Rishu Chaujar***, Physics, DTU

43. Recent Advances of Conducting Polymers and Their Composites for Electrochemical Biosensing Applications, John H.T. Luong, ***6.Tarun Narayan***[1], ***6.Shipra Solanki***[1,2] and ***3.Bansi D. Malhotra***[1], [1]Biotechnology and [2]Chemistry, DTU

44. RISK AND PREVENTING OF COVID-19 IN A COMMERCIAL AIRCRAFT CABIN: AN OVERVIEW, Chanfiou Ahmed Mboreha, Chaharzad Said Abdallah and ***6.Gaurav Kumar***, Mechanical, DTU

45. Scaling Up Solar Manufacturing to Enhance India's Energy Security, Rishabh Jain, ***8.Arjun Dutt*** and Kanika Chawla, Electronics, DTU

46. Software Defect Categorization based on Maintenance Effort and Change Impact using Multinomial Naïve Bayes Algorithm, ***3.Ruchika Malhotra*** and ***8.Madhukar Cherukuri***, CSE, DTU

47. State-of-the-arts Violence Detection using ConvNets, ***7.Aayush Jain*** and ***3.Dinesh Kumar Vishwakarma***, IT, DTU

48. Statistical Analysis of Relation Between Developmental Status of Nations and their Banks' NPA using Hypothesis Testing, ***8.Gangesh Chawla***[1], ***8.Mridul Gupta***[1], ***8.Nimish Sharma***[1], ***8.Keshav Aggarwal***[1], ***3.Ranganath M. Singari***[1] and ***3.Nand Kumar***[2], [1]Mechanical, Production and [2]Humanities, DTU

49. Structural analysis of the chemical vapour deposition grown molybdenum disulphide nanofilms for multifaceted applications, ***8.Sonakshi Saini***, ***8.Sakshi Sharma***, ***8.Maya Khangembam*** and ***3.Vinod Singh***, Applied Physics, DTU

50. Study of Model of Crankshaft and Tools, ***6.SIDDHARTH BOON***, ***3.R.C SINGH***, and ***3.Rajiv Chaudhary***, Mechanical, DTU

51. Tailoring of electrical and optical properties of regenerated silk fibroin films with metal oxides, ***8.Reetu Yadav*** and ***3.Roli Purwar***, Applied Chemistry, DTU

52. TCAD Analysis and Modelling of Gate-Stack Gate All Around Junctionless Silicon NWFET Based Bio-Sensor for Biomedical Application, ***6.Mekonnen Getnet Yirak*** and ***3.Rishu Chaujar***, Applied Physics, DTU

53. Temperature investigation of a Novel 3nm TF-Bulk FinFET for Improved Performance, ***8.Mridul Prakash Kashyap*** and ***3.Rishu Chaujar***, Applied Physics, DTU

54. The Effect of Gate Stack and High-κ Spacer on Device Performance of a Junctionless GAA FinFET, ***7.Bhavya Kumar***, Ajay Kumar and ***3.Rishu Chaujar***, Physics, DTU

55. Twitter Sentiment Analysis Using Lexical or Rule Based Approach: A Case Study, ***7.Sheresh Zahoor*** and ***3.Rajesh Rohilla***, Electronics, DTU

56. UV-shielding antimicrobial zein films blended with essential oils for active food packaging, ***6.Manjot Kaur*** and ***3.Deenan Santhiya***, Applied Chemistry, DTU

57. Viscous Ricci dark energy model with matter creation: Exact solution and observational tests, ***6.AJAY KUMAR*** and ***3.C P SINGH***, Applied Mathematics, DTU

58. Vyaktitv: A Multimodal Peer-to-Peer Hindi Conversations based Dataset for Personality Assessment, Shahid Nawaz Khan, ***8.Maitree Leekhay***, Jainendra Shukla and Rajiv Ratn Shah, DTU

59. Watermarking of Multi-mode Counter Circuit using Hardware Efficient Algorithm, ***3.Jeebananda Panda***, ***8.1.Divant Jain***, ***3.Lavi Tanwar*** and Sunil Kumar, Electronics, 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*

This paper has been submitted to IEEE for publication. Copyright may be transferred without notice.

6G for Bridging the Digital Divide: Wireless Connectivity to Remote Areas

Abdelaali Chaoub, *Senior Member, IEEE*, Marco Giordani, *Member, IEEE*, Brejesh Lall, *Member, IEEE*, Vimal Bhatia, *Member, IEEE*, Adrian Kliks, *Senior Member, IEEE*, Luciano Mendes, Khaled Rabie, *Senior Member, IEEE*, Harri Saarnisaari, *Senior Member, IEEE*, Amit Singhal, *Member, IEEE*, Nan Zhang, *Member, IEEE*, Sudhir Dixit, *Life Fellow, IEEE*

arXiv:2009.04175v1 [cs.NI] 9 Sep 2020

***Abstract*—In telecommunications, network sustainability as a requirement is closely related to equitably serving the population residing at locations that can most appropriately be described as remote. The first four generations of mobile communication ignored the remote connectivity requirements, and the fifth generation is addressing it as an afterthought. However, sustainability and its social impact are being positioned as key drivers of sixth generation's (6G) standardization activities. In particular, there has been a conscious attempt to understand the demands of remote wireless connectivity, which has led to a better understanding of the challenges that lie ahead. In this perspective, this article overviews the key challenges associated with constraints on network design and deployment to be addressed for providing broadband connectivity to rural areas, and proposes novel approaches and solutions for bridging the digital divide in those regions.**

***Index Terms*—6G; remote areas; digital divide; network sustainability; wireless networks.**

I. Introduction

In 2018, 55% of the global population lived in urban areas. Further, 67% of the total World's population had a mobile subscription, but only 3.9 billion people were using Internet, leaving 3.7 billion unconnected, with many of those living in remote or rural areas [1]. People in these regions are not part of the information era and this digital segregation imposes several restrictions to their daily lives. Children growing up without access to the latest communication technologies and online learning tools are unlikely to be competitive in the job and commercial markets. Unreliable Internet connection also hinders people from remote areas to benefit from online commerce and engage in the digital world, thereby compounding already existing social and economic inequalities.

Abdelaali Chaoub is with the National Institute of Posts and Telecommunications (INPT), Morocco (email: chaoub.abdelaali@gmail.com). Marco Giordani is with the Department of Information Engineering, University of Padova, Padova, Italy (email: giordani@dei.unipd.it). Brejesh Lall is with the Indian Institute of Technology Delhi, India (email: brejesh@ee.iitd.ac.in). Vimal Bhatia is with the Indian Institute of Technology Indore, India (email: vbhatia@iiti.ac.in). Adrian Kliks is with the Poznan University of Technology's Institute of Radiocommunications, Poland (email: adrian.kliks@put.poznan.pl). Luciano Mendes is with the National Institute of Telecommunications (Inatel), Brazil (email: luciano@inatel.br). Khaled Rabie is with the Manchester Metropolitan University, UK (email: k.rabie@mmu.ac.uk). Harri Saarnisaari is with the University of Oulu, Finland (email: harri.saarnisaari@oulu.fi). Amit Singhal is with the Bennett University, India (email: singhalamit.iitd@gmail.com). Nan Zhang is with the Department of Algorithms, ZTE Corporation (email: zhang.nan152@zte.com.cn). Sudhir Dixit is with the Basic Internet Foundation and University of Oulu (email: sudhir.dixit@ieee.org).

However, rural areas are now becoming more and more attractive as the new coronavirus (COVID-19) pandemic has shown, since it has reshaped our living preferences and pushed many people to work remotely from wherever makes them most comfortable [2]. Such agglomerations where people live and work are referred to as "oases" in this paper. Wireless connectivity in rural areas is expected to have a significant economic impact too. Hence, the use of technology in farms and mines will increase the productivity and open new opportunities for local communities. Technology will also provide better education, higher quality entertainment, increased digital social engagement, enhanced business opportunities, higher income, and efficient health systems to those living in the most remote zones.

Despite these premises, advances in the communication standards towards provisioning of wireless broadband connectivity to remote regions have been, so far, relegated to the very bottom, if not entirely ignored. The fundamental challenges are low return on investment, inaccessibility that hinders deployment and regular maintenance of network infrastructures, and lack of favorable spectrum and critical infrastructure such as backhaul and power grid, respectively. In these regards, despite being in its initial stages, the 6th generation (6G) of wireless networks is building upon the leftover from the previous generations [3], and will be developed by taking into account the peculiarities of the remote and rural sector, with the objective of providing connectivity for all and reach digital inclusion [4]. Specifically, the research community should ensure that this critical market segment is not overlooked in favor of the more appealing research areas such as artificial intelligence (AI), machine learning (ML), terahertz communications, 3D augmented reality (AR)/virtual reality (VR), and haptics.

Boosting remote connectivity can start by addressing spectrum availability issues. Licenced spectrum in sub-1 GHz, in fact, is a cumbersome and costly resource, and may require new frequency reuse strategy in remote regions because of their unique requirements. Utilization of locally unexploited frequencies and unlicensed bands judiciously may help in reducing the overall cost, thereby making remote connectivity a viable business opportunity. Advanced horizontal and vertical spectrum sharing models, along with enhanced co-existence schemes, are two other powerful solutions to improve signal reach in these areas. Innovative business and regulator models may be suitable, to encourage new players, such as community-based micro-operators, to build and operate the local networks. Local, flexible and pluralistic spectrum

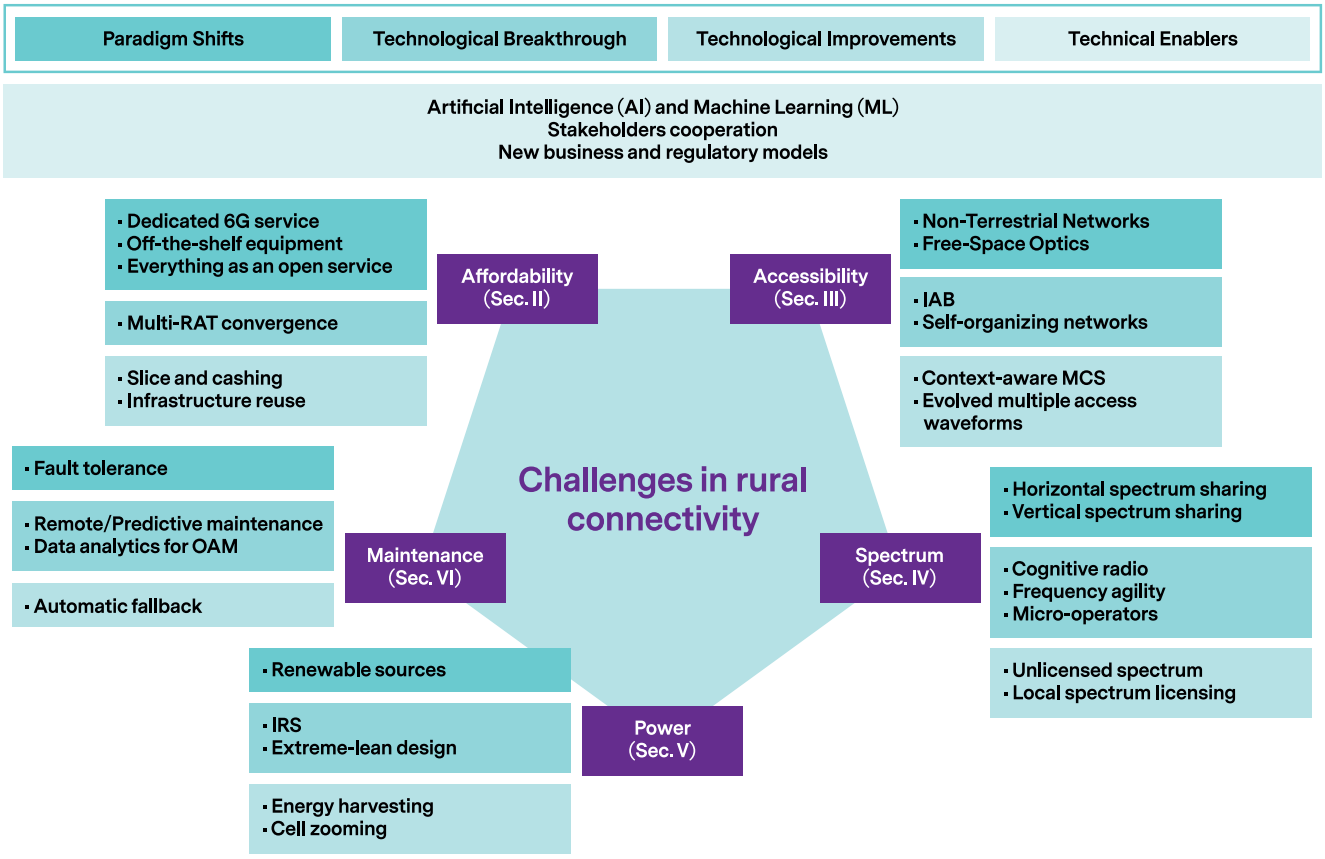


Fig. 1: A summary of the challenges for providing wireless connectivity to remote areas with proposed solutions.

licensing could be the way forward to boost the remote market.

Another issue is that remote areas may not have ample connectivity to the power sources. Hence, it is imperative that 6G solutions for remote areas are designed as self-reliant in terms of their power/energy requirements, and/or with the capability to scavenge from the surrounding, possibly scarce, resources. Governments can assuage this situation to an extent by making it attractive for the profit-wary service providers to deploy solutions in remote areas. Revised government policies and appropriate business models should be parallelly explored as they have direct implications on the technology requirements. Environmentally-friendly thinking should also be included throughout the chain of energy consumed from mining to manufacturing and recycling. Moreover, the abundant renewable sources need to be integrated into power systems at all scales for sustainable energy provisioning.

Remote maintenance of network infrastructures and incorporation of some degrees of self-healing capability is also very important since it might be difficult to access remote areas due to a difficult terrain, harsh weather, or lack of transport-connectivity. Suitable specifications for fault tolerance and fallback mechanisms need therefore to be incorporated.

Based on the above introduction, the objective of this article is two-fold: (i) highlight the challenges that hinder progress in the development and deployment of solutions for catering to the remote areas, and (ii) suggest novel approaches to address those challenges. In particular, the paper targets the 6G mobile standard, such that these important issues are considered into the design process from the very beginning. We deliberately skip a detailed literature survey, because a clear and comprehensive review is provided in [4]. We focus, rather, on discussing the requirements and the corresponding challenges, and proposing novel approaches to address some of those issues.

A summary of these challenges and possible solutions is shown in Fig 1. The rest of the article is organized as follows. Sec. II discusses the question of how future 6G can deliver affordable connectivity to remote users. Sec. III provides a range of promising technical solutions capable of facilitating access to broadband connectivity in remote locations. Sec. IV promotes the use of a variety of dynamic spectrum access schemes and suggests how they can evolve to meet the surging needs in the unconnected areas. Sec. V presents approaches for integrating infrastructure sharing, renewable sources, and emerging energy-efficient technologies to boost optimal and environmentally friendly power provision. Sec. VI presents innovative ways to simplify maintenance operations in hard-to-reach zones. Finally, the conclusions are summarized in Sec. VII.

II. Affordable Service Provisioning in Remote Segments

One of the biggest impediments to connecting the unconnected part of the world is the high costs involved and the prevailing low income of the target population. Fortunately,

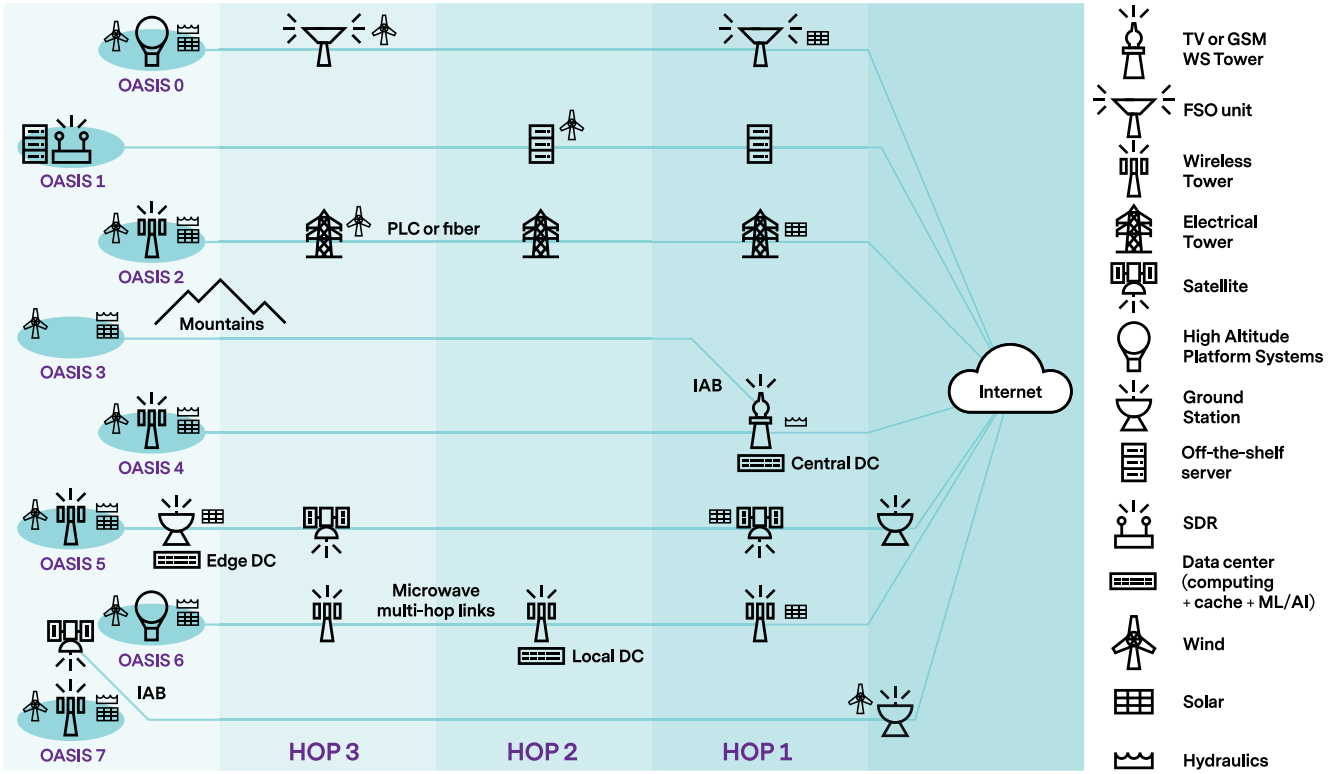


Fig. 2: Backhaul connectivity in rural areas.

there are many affordable emerging alternatives in 6G which may bring new possibilities, as enumerated in this section.

Dedicated remote-centred connectivity layer. Besides 5G's typical service pillars (i.e., eMBB, ULRRC, and mMTC), 6G should introduce a fourth service grade with basic connectivity target key performance indicators (KPIs). However, this remote mode cannot be just a plain version of the urban 6G, since it has to be tailored to the specificities of the remote sector. Some KPIs relevant to remote connectivity scenarios like coverage and cost-effectiveness need to be expanded, whereas the new service class needs more relaxed constraints in terms of some conventional 5G performance metrics like throughput and latency. This novel service class should have its dedicated slice and endowed with specific and moderate levels of edge and caching capabilities: the involved data can then be processed on edge, local or central data centers for better scalability, as illustrated in Fig. 2. Accordingly, such connectivity services can be charged at reduced prices.

Multiple radio access technologies (RATs) interworking. Local access in remote areas can be designed to aggregate multiple and heterogeneous RATs. Remote streams can then be split over one or more RATs, thus allowing flexibility and providing the highest performance possible at minimal cost in everyday life and work. At the same time, digitalization in remote areas calls for large coverage solutions (e.g., TV or GSM white spaces (WSs)) to increase the number of users within a base station and helps reduce the network deployment and management costs, albeit at some performance trade-offs.

Radio frequency (RF) solutions can be complemented by the emerging optical wireless communications (OWCs). In particular, short range visible light communications (VLCs) category operating over the visible spectrum can boost the throughput in indoor, fronthaul and underwater environments (see Fig. 3) while serving the intuitive goal of illumination making it a cost-efficient technology.

Low-cost networking and end-user devices. One way to reduce cost is the exploitation of legacy infrastructure. TV stations can be shared with mobile network operators (MNOs) to provide both tower and electricity. The latest developments in wireless communications can be applied in outdoor power line communication (PLC) to provide high data rate connectivity over the high and medium voltages power lines, increasing the capability of the backhaul networks in remote areas. Existing base stations and the already-installed fibers alongside roads or embedded inside electrical cables can also serve as a backhaul solution for connectivity in rural regions. End-user devices and modems should also be affordable and usable everywhere, i.e., when people move or travel to different places under harsh conditions. Therefore, the possibility to use off-the-shelf equipment at both the user's and network's sides is important and integration with appropriate software stacks is welcome to reduce capital and operational expenditures (capex and opex).

Adoption of open, virtualized and cloud-native solutions. The remote infrastructure is likely to be deployed by small Internet service providers (ISPs) and the cost of specialized hardware equipment is an issue to be overcome. Open source approaches allow MNOs to choose common hardware from any vendor and implement the radio access network (RAN) and core functionalities using software defined radio (SDR)

and software defined networking (SDN) frameworks. Moreover, virtualized and cloudified network functions may reduce infrastructure, maintenance and upgrade costs [5]. These solutions are especially interesting for new players building the remote network from scratch, to foster the inter-operability and cost-effectiveness of hardware and software. However, this field still requires further research and development work before commercial deployment.

III. Improving Service Accessibility in Remote Areas

In order to provide long-lived broadband connectivity, a minimum service quality must be continuously guaranteed. In this perspective, this section reviews potential solutions to promote resilient service accessibility in rural areas.

Multi-hop network elasticity. The access network has, over generations, become multi-hop to provide flexibility in the architecture design, despite some increase in complexity. Given the typical geographic, topographic, and demographic constraints of present scenarios, performance levels (e.g., coverage, latency, and bandwidth) of individual hops can be made adaptive. The idea is to extend performance elasticity beyond air-interface to include other hops in the RAN. The same approach can be brought to backhaul connections (see Fig. 2). Similarly, rural cell boundaries experiencing poor coverage can reap the elasticity benefits through the use of device-to-device (D2D) communications as depicted in Fig. 3.

Network protocols should be extended to include static- (e.g., location-based) besides temporal-quality adaptation to handle variations in channel quality over time.

Wireless backhaul solutions. Service accessibility in rural areas involves prohibitive deployment expenditures for network operators and requires high-capacity backhaul connections for several different use cases. Fig. 2 provides a comprehensive overview of potential backhaul solutions envisioned in this paper to promote remote connectivity. On one side, laying more fiber links substantially boost broadband access in those areas, but at the expense of increased costs. PLC connections, on the other side, provide ease of reach at lower costs making use of ubiquitous wired infrastructures as a physical medium for data transmission, but some inherent challenges related to harsh channel conditions and connected loads are still to be overcome. Fig. 2 illustrates also how, even though the use of conventional microwave and satellite links can fulfill the performance requirements of hard-to-reach zones, emerging long-range wireless technologies, such as TV and GSM WS systems, are capable of delivering the intended service over longer distances with less power while penetrating through difficult terrain like mountains and lakes.

Another recent trend is building efficient cost-effective backhaul links using software-defined technology embedded into off-the-shelf multi-vendor hardware to connect the unconnected remote communities (e.g., Oasis 1 in Fig. 2). Recently, the research community has also investigated integrated access and backhaul (IAB) as a solution to replace fiber-like infrastructures with self-configuring easier-to-deploy relays operating through wireless backhaul using part of the access link radio resources [6]. For example, the TV WS tower in Fig. 2 may use the TV spectrum holes to provide both access to Oasis 3 and connection to the backhaul link for Oasis 4. IAB has lower complexity as compared to fiber-like networks and facilitates site installation in rural areas where cable buildout is difficult and costly. The potential of the IAB paradigm is magnified when wireless backhaul is realized at millimeter waves (mmWaves), thus exploiting a much larger bandwidth than in sub-6-GHz systems. Moreover, mmWave IAB enables multiplexing the access and backhaul data within the same bands, thereby removing the need for additional hardware and/or spectrum license costs.

Nowadays, free space optical (FSO) links are being considered as a powerful full-duplex and license-free alternative to increase network footprint in isolated areas with challenging terrains. However, FSO units are very sensitive to optical misalignment. For instance, the HOP1 FSO unit depicted in Fig. 2 should be permanently and perfectly aligned with the FSO unit installed in the HOP3 location. In-depth research in spherical receivers and beam scanning is hence needed to improve the capability of intercepting laser lights emanating from multiple angles.

Physical-layer solutions for front/mid/backhaul. Even though wireless backhauling can reduce deployment costs, service accessibility in rural regions still requires a minimum number of fiber infrastructures to be already deployed. Fiber capacity can hence be increased if existing wavelength division multiplexing networks are migrated to elastic optical networks (EONs) by technology upgradation at nodes; the outdated technology of urban regions may then be reused to establish connectivity in under-served rural regions without significant investment.

Besides backhaul, midhaul and fronthaul should also be improved by AI/ML-based solutions providing cognitive capabilities for prudent use of available licensed and unlicensed spectrum [7]. This is especially useful in remote areas where the sparse distribution of users may result in spectrum holes. The unlicensed spectrum, in particular, can provide significant cost-savings for service delivery and improve network elasticity. New possibilities including evolved multiple access schemes and waveforms, like non-orthogonal multiple access (NOMA) for mMTC, should be investigated; this technology is particularly interesting for Internet of things (IoT) services where some sensors are close to and some far away from a base station [8]. AI/ML can be also exploited to control physical and link layers for smooth and context-aware modulation and coding schemes (MCSs) transitions, even though this approach would need to be lightweight to reduce cost and maintenance, and optimized for the intended market segment.

Non-terrestrial network solutions. Network densification in rural areas is complicated by the heterogeneous terrain that may be encountered when installing fibers between cellular stations. To solve this issue, 6G envisions the deployment of non-terrestrial networks (NTNs) where air/spaceborne platforms like unmanned aerial vehicles (UAVs), high altitude platform stations (HAPSs), and satellites, provide ubiquitous global connectivity when terrestrial infrastructures are unavailable [9]. Potential beneficiaries of this trend are shown in

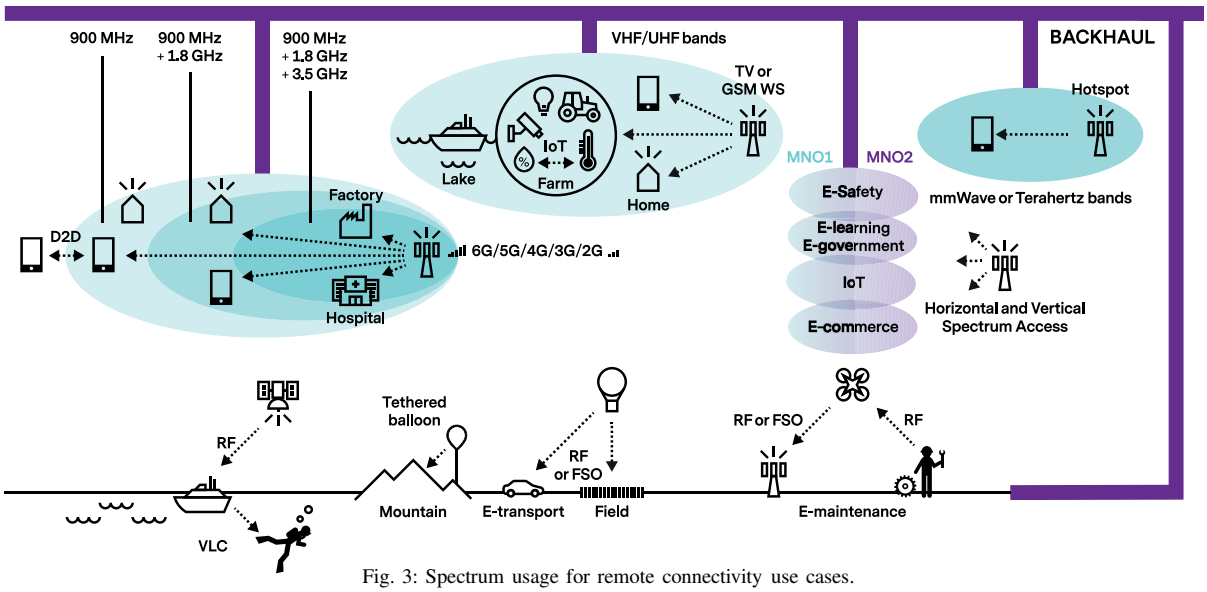


Fig. 3: Spectrum usage for remote connectivity use cases.

Fig. 3, including inter-regional transport, farmlands, ships, mountainous areas, and remote maintenance facilities. The evolution towards NTNs will be favored by architectural advancements in the aerial/space industry (e.g., through solid-state lithium batteries and Gallium Nitride technologies), new spectrum developments (e.g., by transitioning to mmWave and optical bands), and novel antenna designs (e.g., through reconfigurable phased/inflatable/fractal antennas realized with metasurface material). Despite these premises, however, there are still various challenges that need to be addressed, including those related to latency and coverage constraints. NTNs can also provide remote-ready, low-cost (yet robust), and long-range backhaul solutions for terrestrial devices with no wired backhaul.

Self-organizing networks (SONs). To explicitly address the problem of network outages (e.g., due to backhaul failure), which are very common in remote locations, 6G should transition towards SONs implementing network slicing, dynamic spectrum management, edge computing, and zero-touch automation functionalities. This approach provides extra degrees of freedom for combating service interruptions, and improves network robustness. In this context, AI/ML can help both the radio access and backhaul networks to self-organize and self-configure themselves, e.g., to discover each other, coordinate, and manage the signaling and data traffic.

IV. Towards a Flexible Use of Spectrum in Remote Areas

We now present some promising solutions to address spectrum availability issues, which currently pose a serious impediment to broadband connectivity in remote areas.

Leveraging cognitive radio networks. One of the major barriers for network deployment in rural areas is spectrum licensing, since participation in spectrum auction is typically difficult, from an economic point of view, for small ISPs. In this perspective, new licensing schemes can prosper the cognitive radio approach, allowing local ISPs to deploy networks in areas where large operators are not interested in providing their service [10]. Spectrum awareness mechanisms, e.g., geolocation database and spectrum sensing, can be used to inform network providers about vacant spectrum in a given area, as well as providing protection against unauthorized transmissions and unpredictable propagation conditions. For instance, Fig. 3 shows how TV and GSM WS towers can expand the connectivity beyond the rural households to reach more distant locations like farms and wilderness areas.

Spectrum co-existence. Sub-6 GHz frequencies remain critical for remote connectivity thanks to their favourable propagation properties and wide reach. In these crowded bands, spectrum re-farming and inter/intra-operator spectrum sharing can considerably increase spectrum availability [10]. Nevertheless, coverage gaps and low throughput in the legacy bands call for advanced multi-connectivity schemes to combine frequencies above and below 6 GHz. Using advanced carrier aggregation techniques in 6G systems, the resource scheduling unit can choose the optimal frequency combination(s) according to service requirements, device capabilities, and network conditions. The proposed model offers a scalable bandwidth that maintains service continuity in case of connectivity loss in those spectrum bands that are more sensitive to surrounding relief, atmospheric effects, and water absorption: for example Fig. 3 illustrates a scenario in which vital facilities in rural communities enjoy permanent connectivity using the lower bands in case of communication failure on the higher bands. Likewise, multi-connectivity provides diversity, improved system resilience, and situation awareness by establishing multiple links from separate sources to one destination. This aggregation can be achieved at various protocol and/or architecture levels ranging from the radio link up to the core network, allowing effortless deployments of elastic networks in areas difficult to access.

Utilizing unlicensed bands. A combination of licensed and unlicensed bands has been acknowledged by many stan-

dardization organizations to improve network throughput and capacity in unserved/under-served rural areas, as depicted in Fig. 3. While the FCC has recently released 1.2 GHz in the precious 6 GHz bands to expand the unlicensed spectrum, the huge bandwidth available at millimeter- and terahertz-wave bands will further support uplink and downlink split, in addition to hybrid spectrum sharing solutions that can adaptively orchestrate network operations in the licensed and unlicensed bands. High frequencies require line of sight (LOS) for proper communication, complicating harmonious operation with lower bands. Accordingly, time-frequency synchronization, as well as control procedures and listening mechanisms, like listen-before-talk (LBT), need to evolve towards more cooperative and distributed protocols to avoid misleading spectrum occupancy. The management of uncoordinated competing users in unlicensed bands will emerge as important issue, and it needs to be addressed in 6G networks.

Regional licenses and micro-operators. Deployment of terrestrial networks for remote areas is challenging due to terrain, lack of infrastructure and personnel. Network operators would then rather roam their services from telecommunication providers already operating in those areas than building their own infrastructure. However, such an approach may entail the need for advanced horizontal (between operators of the same priorities) and vertical (when stakeholders of various priorities coexist) spectrum/infrastructure sharing frameworks. Solutions like license shared access (LSA, in Europe) and spectrum access system (SAS, in the US) are mature examples of such an approach with two-tiers and three-tiers of users, respectively. This can evolve to include n-tiers of users belonging to m different MNOs. An example of a four-tiered access is provided in Fig. 3. From the top, we find the E-safety services with the highest priority, a tier-2 layer devoted to E-learning sessions and E-government transactions, a middle-priority tier-3 layer for IoT use cases that generate sporadic traffic, and a final lower-priority tier-4 layer that uses the remainder of the available spectrum (e.g., for E-commerce services). Such solutions, however, need to be supported by innovative business and regulatory models to motivate new market entrants (e.g., micro-operators, which are responsible for last-mile service delivery and infrastructure management) to offer competitive and affordable services in remote zones [11].

V. Powering in a Green and Efficient Way

Power supply is among the highest expenses of MNOs and a major bottleneck for ensuring reliable connectivity in remote areas. MNOs' profitability and reliable powering can be improved following (a combination of) these solutions, as summarized in Fig. 4.

Infrastructure sharing. Local communication/power operators, as well as various stockholders such as companies, manufacturers, governmental authorities and standardization bodies, should build an integrated design which entails a joint network development process right from the installation phase. In particular, the different players should cooperate to avoid deploying several power plants for different use cases, thus saving precious (already limited) economic resources for other types of expenses.

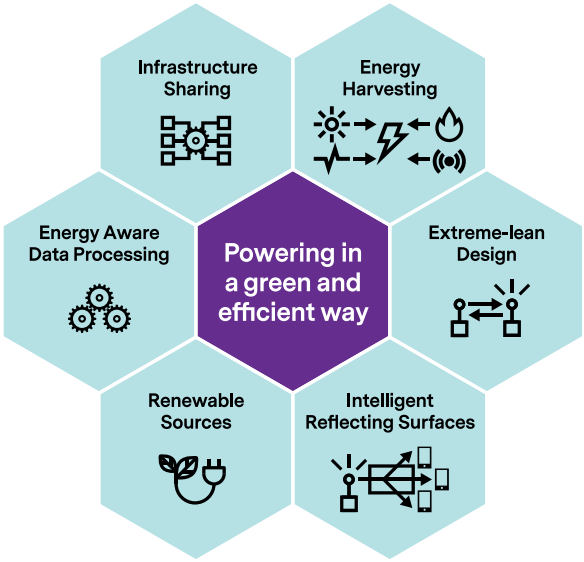


Fig. 4: Key enablers for efficient and green powering of rural areas.

Efficient and optimal energy usage. The 6G remote area solutions should be energy efficient and allow base and relay stations to minimize power consumption while guaranteeing affordable yet sufficient service for residents [12]. In particular, energy efficiency should target IoT sensors' design and deployment, since the increasing use of a massive number of IoT devices, e.g., to boost farming and other activities such as environmental monitoring, is expected to significantly increase in the near future.

At the moment, these efforts have been made after the standardization work was completed, but 6G should include efficient use of energy during the standardization process itself. Techniques like cell zooming relying on power control and adaptive coverage can be reused at various network levels for flexible, energy-saving front/mid/backhaul layouts. AI/ML techniques can be very helpful in these scenarios. For example, the traffic load statistics on each node can be monitored to choose the optimal cell sleeping and on/off switching strategies to deliver increased power efficiency in all the involved steps of communication.

Technological breakthroughs. In addition to obvious energy sources such as solar, wind, and hydraulics, energy harvesting through the ambient resources (e.g., electromagnetic signals, vibration, movement, sound, and heat) could provide a viable efficient solution by enabling energy-constrained nodes to scavenge the energy while simultaneous wireless information and power transfer [13].

Another recent advancement promoting energy-efficient wireless operations is the use of intelligent reflecting surfaces (IRSs), equipped with a large number of passive elements smartly coordinated to reflect any incident signal to its intended destination without necessitating any RF chain. Although still in its infancy, this technology offers significant advantages in making the propagation conditions in harsh remote areas more favorable with substantial energy savings [14].

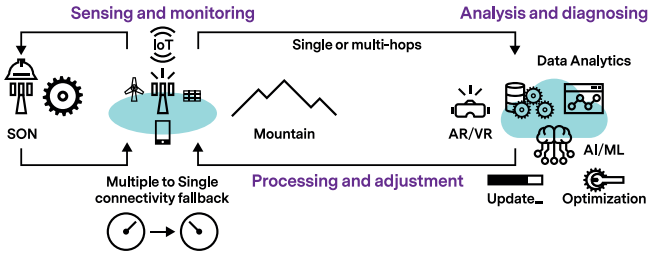


Fig. 5: A workflow for cost-efficient maintenance operations in remote areas.

VI. Intelligent and Affordable Maintenance

Operations, administration and management (OAM) functionalities and dedicated maintenance for each network component are of paramount importance to overall system performance and user experience in traditional commercial 4G/5G networks. This comes at the expense of complicated and costly tasks, especially in hard-to-reach areas. In this section we present innovative ideas to enable intelligent and cost-effective maintenance in 6G network deployed in rural regions.

Network status sensing and diagnosing. Traditionally, the OAM system is adopted for network status monitoring with a major drawback, i.e, manual post-processing and reporting time delay due to huge amounts of gathered data. To enable an intelligent and predictive maintenance, network diagnostics relying on AI-based techniques is advised [15]. With the development of edge computing technologies, multi-level sensing can be employed to achieve near-real time processing and multi-dimensional information collection within a tolerable reporting interval. For instance, processing operations related to short-term network status could be mostly done at the edge node to ensure fast access to this vital information in rural zones.

Network layout planning and maintenance. As mentioned in the previous sections, the network in remote areas is mainly composed of cost-effective nodes along the path from the access to the core parts (e.g., radio, centralized and distributed units, IAB-donors and relays) that need to be organized in either single or multiple hops. In this situation, the whole system will be harmed if one of these nodes experiences an accidental failure. To enhance the resilience of such networks, more flexible and intelligent network layout maintenance is required. More precisely, using evolved techniques such as SONs (see Sec. III), the link among each couple of nodes within the network can be permanently controlled and dynamically substituted or restored in case of an outage (see Fig. 5). Additionally, since a big part of the next generation mobile network is virtualized, appropriate tools or even a dedicated server may be needed for automatic software updates monitoring, periodic backups and scheduled maintenance to avoid or at least minimize the need for on-site intervention in those remote facilities. Automatic fallback mechanisms can also be scheduled to downgrade the connectivity to another technology under bad network conditions, e.g., by implementing appropriate multi-connectivity schemes, as described in Sec. IV.

Network performance optimization. Network optimization in rural areas should take into account remote-specific requirements and constraints. For example, access to the edge resources, which are finite and costly and can be rapidly exhausted, should be optimized taking into consideration the intended services, terminal capabilities, and charging policy of the network and its operator(s).

A summary of the maintenance life cycle in remote and rural areas is shown in Fig. 5. In particular, after intelligently building and processing relevant system information data sets, maintenance and repair activities (e.g. system updates or operational parameters optimization) can be performed remotely and safely using 3D virtual environments such as AR and VR.

VII. Conclusions

The problem of providing connectivity to rural areas will be a pillar of future 6G standardization activities. In this article we discuss the challenges and possible approaches to addressing the needs of the remote areas. It is argued that such service should be optimized for providing a minimum fallback capability, while still providing full support for spatio-temporal service scalability and graceful quality degradation. We also give insights on the constraints on network design and deployment for rural connectivity solutions. We claim that optimally integrating NTN and FSO technologies along the path from the end-point to the core element using open software built on the top of off-the-shelf hardware can provide low-cost broadband solutions in extremely harsh and inaccessible environments, and can be the next disruptive technology for 6G remote connectivity. Integration of outdated technology should also be provisioned so that they may be innovatively used to service the remote areas. Such provisions should extend to integrate open and off-the-shelf solutions to fully benefit from cost advantage gains. Spectrum, regulatory, and standardization issues are also discussed because of their importance to achieve the goal of remote area connectivity. It is fair to say that including remote connectivity requirements in the 6G standardization process will lead to a more balanced and universal social as well as digital equality.

References

[1] H. Saarnisaari, S. Dixit, S. Alouini, A. Chaoub, M. Giordani, A. Kliks, M. Matinmikko-Blue, and N. Zhang, " 6G White Paper on Connectivity for Remote Areas," *6G Research Visions, No. 5, University of Oulu*, 2020.

[2] J. Phillipson, M. Gorton, R. Turner, M. Shucksmith, K. Aitken-McDermott, F. Areal, P. Cowie, C. Hubbard, S. Maioli, R. McAreavey *et al.*, "The COVID-19 Pandemic and Its Implications for Rural Economies," *Sustainability*, vol. 12, no. 10, p. 3973, May 2020.

[3] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, "Toward 6G Networks: Use Cases and Technologies," *IEEE Communications Magazine*, vol. 58, no. 3, pp. 55–61, March 2020.

[4] E. Yaacoub and M. Alouini, "A Key 6G Challenge and Opportunity – Connecting the Base of the Pyramid: A Survey on Rural Connectivity," *Proceedings of the IEEE*, vol. 108, no. 4, pp. 533–582, Mar. 2020.

[5] L. Gavrilovska, V. Rakovic, and D. Denkovski, "From Cloud RAN to Open RAN," *Wireless Personal Communications*, pp. 1–17, March 2020.

[6] M. Polese, M. Giordani, T. Zugno, A. Roy, S. Goyal, D. Castor, and M. Zorzi, "Integrated Access and Backhaul in 5G mmWave Networks: Potential and Challenges," *IEEE Communications Magazine*, vol. 58, no. 3, pp. 62–68, March 2020.

[7] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y.-J. A. Zhang, "The roadmap to 6G: AI empowered wireless networks," *IEEE Communications Magazine*, vol. 57, no. 8, pp. 84–90, Aug 2019.

[8] Z. Hu, L. Xu, L. Cao, S. Liu, Z. Luo, J. Wang, X. Li, and L. Wang, "Application of non-orthogonal multiple access in wireless sensor networks for smart agriculture," *IEEE Access*, vol. 7, pp. 87 582–87 592, Jun 2019.
[9] N. Cheng, W. Quan, W. Shi, H. Wu, Q. Ye, H. Zhou, W. Zhuang, X. S. Shen, and B. Bai, "A comprehensive simulation platform for space-air-ground integrated network," *IEEE Wireless Communications*, vol. 27, no. 1, pp. 178–185, Feb 2020.
[10] W. S. H. M. W. Ahmad, N. A. M. Radzi, F. Samidi, A. Ismail, F. Abdullah, M. Z. Jamaludin, and M. Zakaria, "5G technology: Towards dynamic spectrum sharing using cognitive radio networks," *IEEE Access*, vol. 8, pp. 14 460–14 488, Jan 2020.
[11] P. Ahokangas, M. Matinmikko-Blue, S. Yrjölä, V. Seppänen, H. Hämmäinen, R. Jurva, and M. Latva-aho, "Business models for local 5G micro operators," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 3, pp. 730–740, Mar 2019.
[12] GSMA Connected Society, "Closing the Coverage Gap – How Innovation Can Drive Rural Connectivity," *White Paper*, 2019.
[13] M. Gholikhani, H. Roshani, S. Dessouky, and A. Papagiannakis, "A critical review of roadway energy harvesting technologies," *Applied Energy*, vol. 261, p. 114388, Mar 2020.
[14] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Communications Magazine*, vol. 58, no. 1, pp. 106–112, Nov 2019.
[15] M. Huang, Z. Liu, and Y. Tao, "Mechanical fault diagnosis and prediction in IoT based on multi-source sensing data fusion," *Simulation Modelling Practice and Theory*, vol. 102, p. 101981, July 2020.

Abdelaali Chaoub [SM] is an Associate Professor, working at the National Institute of Posts and Telecommunications (INPT) at Rabat (Morocco) since 2015. His research interests are related to spectrum sharing for 5G/B5G networks, cognitive radio networks, smart grids, cooperative communications in wireless networks, and multimedia content delivery. He is a paper reviewer for several leading international journals and conferences. He has accumulated intersectoral skills through work experience both in academia and industry as a Senior VoIP solutions Consultant at Alcatel-Lucent (2007–2015).

Marco Giordani [M'20] received his Ph.D. in Information Engineering in 2020 from the University of Padova, Italy, where he is now a postdoctoral researcher and adjunct professor. He visited NYU and TOYOTA Infotechnology Center, Inc., USA. In 2018 he received the "Daniel E. Noble Fellowship Award" from the IEEE Vehicular Technology Society. His research focuses on protocol design for 5G/6G mmWave cellular and vehicular networks.

Brejesh Lall [M] received his Ph.D degree from Indian Institute of Technology Delhi in 1999. Earlier, he received his bachelors and masters degree in Electronics and Communications from Delhi College of Engineering in 1991 and 1992 respectively. He is currently a Professor in the department of Electrical Engineering at Indian Institute of Technology Delhi. Previously, he served in the Digital Signal Processing Group of Hughes Software Systems for 8 years. His research interests lie in the areas of signal processing and machine learning. He has extensively applied signal processing / machine learning techniques to applications in the broad areas of telecommunications and computer vision.

Vimal Bhatia [M'96, SM'12] is currently working as a Professor at Indian Institute of Technology Indore (IIT-I), Indore, India. He received his Ph.D. degree from Institute for Digital Communications at The University of Edinburgh (UoE), UK in 2005. During his Ph.D. studies he also received the IEEE fellowship for collaborative research on at Carleton University, Ottawa, Canada. He has authored/co-authored more than 230 peer-reviewed journals and conferences. His research focuses on optical communication networks, wireless communications, signal processing and in software product developments.

Adrian Kliks [SM] is an assistant professor at Poznan University of Technology's Institute of Radiocommunications, Poland, and he is a cofounder and board member of RIMEDO Labs company. His research interests include new waveforms for wireless systems applying either non-orthogonal or non-contiguous multicarrier schemes, cognitive radio, advanced spectrum management, deployment and resource management in small cells, and network virtualization.

Luciano Mendes received his Ph.D. on Electrical Engineering from the State University of Campinas, Brazil in 2007. Since 2001 he is a professor at the National Institute of Telecommunications (Inatel), Brazil, where he acts as Research Coordinator of the Radiocommunications Reference Center. His main research area is physical layer for future mobile networks.

Khaled Rabie [M'15, SM'20], a Fellow of the U.K. Higher Education Academy, received his Ph.D. degree from the University of Manchester, UK. He is currently an assistant professor at the Manchester Metropolitan University, UK. His primary research focuses on various aspects of the next-generation wireless communication systems. He received the best student paper award at the IEEE ISPLC (TX, USA, 2015) and the IEEE Access Editor of the month award for August 2019. Khaled is an Editor for IEEE Communications Letters.

Harri Saarnisaari [SM] received his Ph.D degree from the University of Oulu in 2000, where he has been with Centre for Wireless Communications since 1994. He is currently a university researcher and his current research interest include remote area connectivity, especially in the Arctic areas.

Amit Singhal [M] received his PhD degree in electrical engineering from the Indian Institute of Technology Delhi in 2016. He is currently working as an Assistant Professor at Bennett University, Greater Noida, India. His research interests include next generation communication systems, Fourier decomposition method, image retrieval and molecular communications.

Nan Zhang [M] was born in Qingyang, China, in 1990. He received the bachelor degree in communication engineering and the Master degree in integrated circuit engineering from Tongji University, Shanghai, China, in July 2012 and March 2015, respectively. He is now a Senior Engineer at the Department of Algorithms, ZTE Corporation and works on the standardization of LTE and NR system. His current research interests are in the field of 5G channel modeling, MIMO, NOMA techniques, satellite/ATG communication and network architecture.

Sudhir Dixit [LF] is a Co-Founder of the Basic Internet Foundation in Oslo, Norway, and heads its US operations. He is also a Docent at University of Oulu. From 2015 to 2017 he was the CEO and Co-Founder of a start-up, Skydoot, Inc. in California. From 2009 to 2015, he was a Distinguished Chief Technologist and CTO of the Communications and Media Services for the Americas Region of Hewlett-Packard Enterprise Services in Palo Alto, CA, and the Director of Hewlett-Packard Labs India in Palo Alto and Bangalore. Before joining HP, he worked for BlackBerry, NSN, Nokia and Verizon. He has published 8 books, over 200 papers and holds 21 US patents. Dixit holds a Ph.D. from the University of Strathclyde, Glasgow, U.K. and an M.B.A. from the Florida Institute of Technology, Melbourne, Florida.

ADV MATH SCI JOURNAL

Advances in Mathematics: Scientific Journal **9** (2020), no.6, 3537–3548
ISSN: 1857-8365 (printed); 1857-8438 (electronic)
https://doi.org/10.37418/amsj.9.6.33 Spec. Issue on RDESTM-2020

A COMPREHENSIVE OUTLINE OF DIVERSE IDS/IPS WITH RESPECT TO HYPERVISORS

A. N. SESHU KUMAR[1], R. KUMAR YADAV, AND N. S. RAGHAVA

ABSTRACT. The wide exponential growth of cloud computing has opened its applications doors in almost all the areas ranging from business, communication, education, medical, scientific technologies and many more. This widespread on the other hand has given enough space for the intruders to carry out their tasks. Virtualization is the backbone of cloud computing concepts. With the help of virtualization , virtual machines are created in cloud. Virtualization cannot be carried out successfully without proper usage of hypervisors. Depending on the kind of application to be carried out on cloud a hypervisor can be either hardware, software or a firmware. Many instances can be carried out on a virtual machine with the help of these hypervisors which are helpful to run either on guest or host operating system. Intruders in order to gain control on cloud, they first try to achieve control over hypervisors. In this paper we discuss in detail about the various attacks which are being carried out on hypervisors and how to defend and offend them.

[1]*corresponding author*

2010 *Mathematics Subject Classification.* 68T99.

Key words and phrases. Anomaly based IDS (A-IDS), Artificial Intelligence (AI), Genetic Algorithms (GA), Hypervisor, Hybrid based IDS (H-IDS), Hypervisor-based Cloud Intrusion Detection System (HCIDS), Intrusion Detection System (IDS), Side-channel attack (SCA), and Signature based IDS (S-IDS), Virtual Machine (VM) and Virtual Machine Introspection (VMI), Visualization.

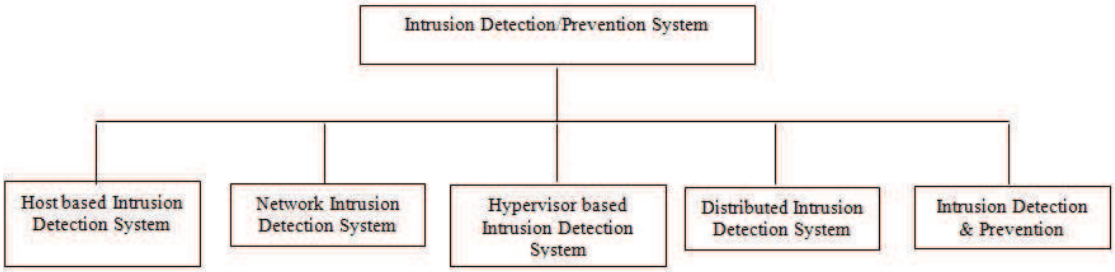


FIGURE 1. Classification of IDS/IPS

1. INTRODUCTION

Information services are deliberately delivered in an exponential order with the help of cloud computing where reliable platforms offer a variety of scalable dynamically provisioning resources for different users who are located at different geographical location. Besides providing computing at low cost and on demand infrastructures, the deliberate extensions of cloud in various fields varying from single user to multiuser, from small and medium business industries to multinational companies. Also, there are various additional features of cloud as irrespective device access, irrespective location access, availability of services throughout the clock, variation of services from public, private, community and hybrid clouds. Due to these numbers of free dimensions cloud has been considered as an opportunity by the intruders to carry out their task. One of the advanced technologies in towards techno world is cloud computing which is helpful for resource sharing through network at various geographically location irrespective of devices these things are performed at low cost. Cloud concepts support various models which are ranging as PaaS, SaaS, IaaS. Virtualization is considered as the backbone of cloud computing. With the help of which resource can be shared among many users at the same time. Entire paper is divided into six main sections where section I deals with the introduction of hypervisors along with certain basic background of cloud concepts, section II ,III and IV mainly deals with a detailed classification of IDS/IPS, Capabilities of IDS/IPS their responses during the attack scenario. Section V deals with an analysis of attacks which are being carried out on hypervisors how they were overcome by using methods along with their limitations. Section V deals with the conclusion

and contribution of future work with two proposals on overcoming the attacks associated with Type II hypervisors.

2. Classification Of IDS/IPS

In a general scenario we can classify various IDS/IPS in the following different categories which are discussed below depending on the position where they are installed.

A. Host Based Intrusion Detection Systems (HIDS)

Here the presence of an agent on the host carries out the required application logs, analyzing system calls, modified file systems; here software agents are present who carry out the task of IDS.

B. Network Based Intrusion Detection System (NIDS)

Unauthorized access of computers, port scans, DoS attacks are detected during monitoring of network by NIDS i.e. any kind of network-based threats can be detected. Intrusions can be detected with a comparison which is carried out with respect to existing attack database to the obtained report by monitoring the network. With the help of Xarp 2 tool using an ARP scanning method [1] attacks with respect to network traffic are traced out. Event collector and sensor [2] gathers information about attack behavior by the data which is sensed while the attack is going.

C. Hypervisor-Based Intrusion Detection System

Hypervisor is a software which when installed at the hypervisor layer between software and hardware then with the help of which VM are created with the help of which intrusion detection is carried out. [3] Proposed a method with the help of HIDS for intrusion detection in virtual environment with the help of VM monitors it observes events, software states, hardware states with the help of UNIX socket commands.

D. Distributed Intrusion Detection Systems (DIDS)

On a large network several intrusions detection systems are combined together for DIDS it is related to a central server with the help of which a continuous monitoring is carried out when any kind of new attack is detected an alert message is generated. An architecture with DIDS is used with a prototype of security monitoring is discussed by [4]. Intrusion is defined as the act

TABLE 1. Showing Comparison Of Different Kinds Of Ids

IDS technique Parameter	Host Based	Distributed	Network
Detection with respect to encrypted data	Does not detect	Detection is not perfect	Moderately used
Cost effectiveness	Less	Less	Moderate
Any OWASP top 10 attack conducted	Limited to two attacks	Limited to three attacks	Limited to four attacks
Accuracy to get the results	Moderate not satisfactory results	Moderate but not accurate results	Reliable
Does it work during high traffic	Works well for light traffic	Works well for moderate traffic	Does not work well for high traffic
Can be used for data segregation and analysis	Data is not perfectly segregated	Moderately segregated	Reliable for segregation
Used for cloud forensics	Moderately used	Moderately used	Moderately used
Does any protection of data methods include	It does not include a full-fledged protection	Provides moderate protection to data	Provides moderate protection to data

carried out by an intruder either to break an information system or to access the data illegally. Internal and External intruders are the two groups which are the classification from intruders. Unauthorized access of systems with the help of assorted techniques of penetrations is referred as external intruders. Unauthorized access of information internally in systems is termed in the category of internal intruders. Exploitation of specific protocols, cracking of passwords, unsecured traffic sniffing, system misconfiguration and software bugs exploiting software. A system which reports and detects intrusions accurately is known as IDS. They are generally operating system specific and operate as an important tool for implementing security policy of an organization. Importance of IDS in tracing out various attacks is being describe clearly by [6].

E. Intrusion Detection and Prevention System (IDPS)

A combination of IDS with firewall is generally used for intrusion prevention here signature with precise traffic rules are used. Passage or blocking of network traffic is decided by IPS depending on the rules which are preconfigured. As soon as an attack is detected by IPS it blocks the attacks and presents further damage. IDPS is obtained when both IDS and IPS are brought together. [5] Describes a network which is efficient for both intrusion detection and protection

where there is no need to install IDS at each node, it automatically generates alerts.

3. Capabilities Of IDS/IPS

A network which generally functions in a usual way deviates from its normal functioning when it is being affected by unauthorized activities either indirect or directly. This also causes serious effects to the persons who are connected to the same network. An IDS plays a vital role in order to trace such kind of activities where not only passive and active activities of the intruder are traced but also one can get an initiation about any kind of mischief which is being carried out with respect to the victims systems or the network associated with him. Following are certain capabilities of IDS.

- Authorized user of a network gets an auto notification or alert during intrusion attacks.
- The friendly environment of IDS allows even a non-IT person also to carryout effective measures during attacks.
- New or existing attacks can be tracked easily which helps to take the counter measures.
- Soon sniffing of any attack, these IDS prevents further loss of data through the network where they are actively implemented. With the help of specific rules set incorporated in the IDS gives more effective results than in normal states. An excellent style manual for science writers is [7].

4. Analysis Of Cloud Based IDS

A flavour of services is enjoyed by the end users delivered by the cloud service providers with respect to the remote servers which are located at certain geographical location. IDS based on cloud are generally classified in the following categories as signature based, Anomaly based, hybrid based.

A. Signature based (S-IDS)

Certain rules (signatures) are developed and maintained in the database. Basing on previous network attacks when similar kind of attacks are being felt in future, these rules are being activated automatically and maximum damage from the network is prevented and when new a attack is encountered its relevant

rules are being framed. The same is updated every time in the database [7]. A cloud-based IDS [8] which is a form or S-IDS is kept at each cloud region which is helpful in detecting any anonymous activities in the network. Beside alarming the fact, it also reduces the loss of data carried out by the intruders. Out of various kind of attacks, DoS and DDoS are traced out easily. The only drawback of this IDS is that it fails to detect new kind of attacks and during overhead computational process it is difficult to trace the attacks. For a given VM maximum protection with respect to DDOS can be carried out with the help of these IDS [9]. Integrated NIDS in an open source cloud environment is also considered as S-IDS [10].

B. Anomaly based IDS in cloud

Here comparison is carried out between a system working under normal condition to a system working during malicious event carried out by intruder on the system. During a long period of time user's legitimate behaviour is compared with various tests in statistical way. Even under the situations where a system is not updated there also unknown attacks are also traced out [11][12]. Its drawback is that it uses a greater number of resources in order to detect unknown attacks with high accuracy [13]. In cloud VM's with respect to every guest OS it is assigned it is a kind of host-based IDS. By the effective usage of Anomaly IDS [14], hyper-call based attacks are traced out for VMM that are affected by hyper call.

C. Hybrid based IDS in cloud

It's the combination of both S-IDS and A-IDS with the help of which both unknown and known attacks are being traced out [15]. These kinds of IDS can be used both in host and network. This kind of IDS is carried out in real time cloud environment even unknown attacks can also be detected with high accuracy with effective usage of ANN applications. [16]. Its drawback is that it must be performed on more and more data for getting correct results of the analyzed data [17]. In cloud if all the required activities are carried out in proper way then it may lead to green computing [18].

5. Overview Of Hypervisor Attacks

Hypervisor attacks can be classified into the following types which are described as below.

A. Hypervisor-based Cloud Intrusion Detection System

System performance during an attack with respect to a cloud instance can be traced successfully using HCIDS [19] where additional installation of software can be overcome. It provides a kind of cloud-based IDS; its performance can be evaluated effectively in real time cloud environment. With this system one can easily detect DoS attacks with respect to cloud instances. The detection of the intrusion system is carried out even outside the virtual machine and independent of the operating systems which are being running on a VM. Drawback of this system is that malicious activities with exact accuracy are missed to be found.

B. Empirical Evaluation of the Hypervisor Scheduling on Side Channel Attacks

A SCA deals with the collection of information related to operations done using cryptography with respect to a computing device. It's a kind of security exploit. [20]. SCA are going to occur when various resources are being shared among multiple VM's. Due to these end users hidden information is being theft which may include as encryption keys it is more effectively carried out by monitoring victim's patterns which are being accessed by him including hardware, CPU cache etc. scheduling schemes of hypervisors can be expressed as an effective way to overcome these kinds of attacks. This information is helpful to design effective hypervisor sharing schemes. The only drawback is that new attacks may not be traced out effectively.

C. Analysis of Potential Hypervisor Attack Vectors

For a reliable software and hardware hypervisor provide a channel. As they share an important part of cloud so they are more prone to attacks. Besides giving host for virtual machine they also share and are helpful in granting access to various resources. With respect to ESXi 5.0 hypervisor platform various kinds of vulnerabilities are discussed [21]. The drawback here is that new attacks cannot be traced out properly with high accuracy.

D. Analysis of Various Virtual Machine Attacks In Cloud Computing

Multi-tenancy is referred as sharing database among many. One way to secure cloud infrastructure is to use the concept of virtualization based on hypervisors. [22] Describes various attacks relevant to security and analyses of the same

with respect to hypervisor security are carried out in an environment of virtualization. A kind of security zone is created by every hypervisor itself. For a virtual machine which is running on a host it is mostly effected by the hypervisor. Within the hypervisor itself there are multiple security zones where it is a bit difficult to trace the attacks at a particular security zone and to provide it relevant measures. With respect to each kind of security zone there are different attacks which are being carried out the intruders. The drawback is that it fails in many issues to deliberately detecting attacks on hypervisors at high level of multi-tenancy.

E. Insider Attacks in Cloud Computing

When attacks made on the cloud are being classified into various groups i.e. whether the attack is insider or outsider attack then it would be easier to take the relevant measures. With the rapid growth of cloud in almost every sector of industry it is prone to number of attacks some of them are carried out from outside the cloud and some of them are carried out from inside the cloud. As our assets relevant to computing may be either inside or outside the cloud when more number or assets are inside the cloud then the risk of the attack is also more. [23] describes the methods to know the nature of insider's attack that would be helpful to take proper measure with respect to the cloud environment. The only drawback of this approach is that for complex attack it is a bit difficulty to classify the attacks into the required subsection which if done properly then necessary measure can be taken to prevent the future kind of such attacks. A detailed study of emerging challenges in the field of cloud computing are discussed by [24] clearly shows that though it constituted a decade for the research which are being carried out in cloud area but still there emerges new threats and security measures which has to be answered.

The below fig. shows the extension of holistic insider's scenario presence in the cloud ecosystem. The detailed description of various threat actors and the location of the insider with in the eco system are explained below.

A. Advanced Persistent Threat

An attacker or intruder present within a network for a long duration to gain the control of all the activities within the system lead to APT.

B. Familial Insiders

A client feels relaxes to himself when he is working in an environment where he feels safe and more secure where he can decrease the security levels.

C. Benigin-insiders coercion

Outsiders who have good relations with insiders in cloud environment may sometime compromise them with their good relations.

D. Malicious Hypervisor:

Without proper knowledge of hypervisor or in genuine usage of hypervisor may also be a reason for data loss or an hotspot for various attacks in clouds.

E. Malicious Clients and CSP:

Sometimes both client and CSP show that they are having trust of each other but they come to know at certain point that they both are lying each other.

6. Conclusion

From this literature study on various attacks carried out on hypervisors in cloud. As hypervisors act as the basic backbone of cloud concepts. Due to various importance's of hypervisors as mentioned in paper so there is an extensive need to protect hypervisor to carry out the sustainable applications on cloud without any halt. We would like to suggest two unique methods which we are going to discuss in our coming research papers which are useful to trace out various attacks on hypervisors. One method is with respect to effective usage of GA in order to find malicious attacks which are being carried out on the hypervisors in the cloud. Second method is to create an effective immune system in order to overcome these attacks thereby enhancing the security levels in cloud concepts with effective usage of certain active artificial intelligence applications.

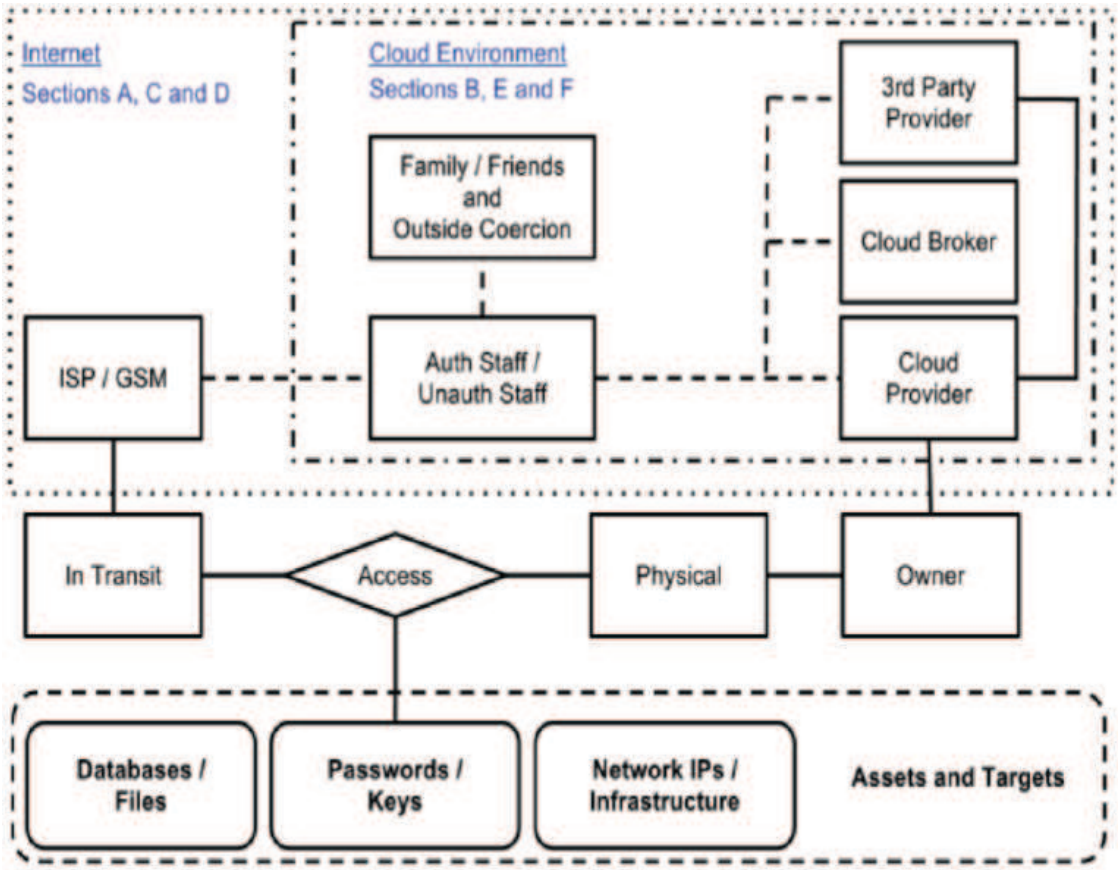


FIGURE 2. Mapping of insider contents to their sub section in cloud eco system

REFERENCES

[1] M.A. HEMAIRY, S. AMIN, Z. TRABELSI: *Towards more sophisticated ARP Spoofing detection/prevention systems in LAN networks*, International conference on the current trends in information technology (CTIT), 2009, 1–6.

[2] S. ROSCHKE, C. FENG, C. MEINEL: *An extensible and virtualization compatible IDS management architecture*, Fifth international conference on information assurance and security, **2** (2009), 130–134.

[3] S. LANIEPCE, M. LACOSTE, M. KASSI-LAHLOU, F. BIGNON: *Engineering Intrusion Prevention Services for IaaS Clouds: The Way of the Hypervisor*, Service Oriented System Engineering (SOSE), IEEE 7th International Symposium on March, 2013.

[4] S. R. SNAPP, J. BRENTANO, G. V. DIAS, T. L. GOAN, T. GRANCE, L. T. HEBERLEIN: *A system for distributed intrusion detection*, IEEE, Compcon, 1991, 170–176.

[5] M. AHMED, R. PAL, H. M. HOSSAIN, M. BIKAS, M. K. HASAN: *NIDS: A Network Based Approach to Intrusion Detection and Prevention*, Computer Science and Information Technology—Spring Conference, 2009, 141–148.

[6] S. K. MOHIDDIN, D. Y. S. BABU: *A Relevance Technical Approach for Screening the Significance of IDS in Cloud Forensics*, IJITEE, 2019, 2278–3075.

[7] J. C. FOSTER: *IDS: Signature versus anomaly detection*, http://search.security.techtarget.com/tip/IDS-Signature-versus-anomaly-detection, 2005.

[8] C. C. LO, C. C. HUANG, J. KU: *A Cooperative Intrusion Detection System Framework for Cloud Computing Networks*, 39th International Conference on Parallel Processing Workshops, 2010, 280–284.

[9] C. MAZZARIELLO, R. BIFULCO, R. CANONICO: *Integrating a Network IDS into an Open Source Cloud Computing Environment*, Sixth International Conference on Information Assurance and Security, 2010, 265–270.

[10] A. PATEL, Q. QASSIM, Z. SHUKOR, J. NOGUEIRA, J. JUNIOR, C. WILLS: *Autonomic Agent-Based Self-Managed Intrusion Detection and Prevention System*, Proceedings of the South African Information Security Multi-Conference (SAISMC), 2010, 223–234.

[11] D. COHEN: *What is a Zero-Day Exploit* http://what-is-what.com/ what is / zero_ day_exploit.html, 2007.

[12] D. MUDZINGWA, R. AGRAWAL: *A study of methodologies used in intrusion detection and prevention systems (IDPS)*, Proceedings of IEEE Southeastcon, 2012, 1–6.

[13] A. V. DASTJERDI, K. A. BAKAR, S. G. H. TABATABAEI: *Distributed Intrusion Detection in Clouds using Mobile Agents*, Third International Conference on Advanced Engineering Computing and Applications in Sciences, 2009, 175–180.

[14] S. BHARADWAJA, W. SUN, M. NIAMAT, F. SHEN: *Collabra: A Xen Hypervisor based Collaborative Intrusion Detection System*, Eighth International Conference on Information Technology: New Generations, 2011, 695–700.

[15] A. PATEL, M. TAGHAVI, K. BAKHTIYARI, J. C. JUNIOR: *An Intrusion Detection and Prevention System in Cloud Computing: A Systematic Overview*, Journal of Network and Computer Applications, **36** (2013), 25–42.

[16] K. VIEIRA, A. SCHULTER, C. B. WESTPHALL, C. M. WESTPHALL: *Intrusion Detection for Grid and Cloud Computing*, IEEE Computer Society, 2010, 38–43.

[17] S. N. DHAGE, B. B. MESHRAM, R. RAWAT, S. PADAWE, M. PAINGAOKAR: *Intrusion Detection System in Cloud Computing Environment*, International Conference and Workshop on Emerging Trends in Technology, 2011, 235–239.

[18] S. K. MOHIDDIN, Y.S. BABU: *Green computing an eco-friendly IT environment for upcoming technologies*, International Journal of Advanced Research in Computer Science, **6** (2015), 28–32.

[19] J. NIKOLAI, Y. WANG: *Hypervisor-based cloud intrusion detection system*, International Conference on Computing, Networking and Communications (ICNC), Honolulu, HI, 2014, 989–993.

[20] L. LIU, A. WANG, W. ZANG, M. YU, S. CHEN: *Empirical Evaluation of the Hypervisor Scheduling on Side Channel Attacks*, IEEE International Conference on Communications (ICC), Kansas City, MO, 2018, 1–6.

[21] L. TURNBULL, J. SHROPSHIRE: *Breakpoints: An analysis of potential hypervisor attack vectors*, Proceedings of IEEE Southeastcon, Jacksonville, FL, 2013, 1–6.

[22] S. ANNAPOORANI, B. SRINIVASAN, G. A. MYLAVATHI: *Analysis of various virtual machine attacks in cloud computing*, 2nd International Conference on Inventive Systems and Control (ICISC), Coimbatore, 2018, 1016–1019.

[23] A. J. DUNCAN, S. CREESE, M. GOLDSMITH: *Insider attacks in cloud computing*, IEEE 11th international conference on trust, security and privacy in computing and communications, 2012, 857–862.

[24] S. K. MOHIDDIN, S. B. YALAVARTHI: *Research Challenges in the Emerging trends of Cloud Computing*, International Journal of Advances in Computer Science and Technology (IJACST), **4**(1) (2015), 4–14.

DEPT. OF COMPUTER SCIENCE ENGG.
DELHI TECHNOLOGICAL UNIVERSITY
DELHI-110042. INDIA
Email address: seshu1203@gmail.com

DEPT. OF COMPUTER SCIENCE AND ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
DELHI-110042. INDIA.
Email address: rkyadav6711@gmail.com

DEPT. OF ELECTRONICS AND COMMUNICATION ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
DELHI-110042. INDIA
Email address: nsraghava@gmail.com

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/344297118

A Framework for the Improvement of Frugal Design Practices

Article · September 2020

DOI: 10.3390/designs4030037

CITATIONS
0

READS
27

3 authors:



Ravindra Singh
Delhi Technological University
8 PUBLICATIONS **8** CITATIONS
SEE PROFILE



Sumedha Seniaray
Delhi Technological University
2 PUBLICATIONS **0** CITATIONS
SEE PROFILE



Prateek Saxena
Cranfield University
25 PUBLICATIONS **93** CITATIONS
SEE PROFILE

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

simulation solar cells with finite element method View project

Membrane Technology View project

All content following this page was uploaded by Prateek Saxena on 18 September 2020.

The user has requested enhancement of the downloaded file.

designs

MDPI

Article

A Framework for the Improvement of Frugal Design Practices

Ravindra Singh [1], Sumedha Seniaray [2] and Prateek Saxena [3,*]

1 Department of Design, Delhi Technological University, Delhi 110042, India; ravindra@dtu.ac.in
2 Department of Applied Mathematics, Delhi Technological University, Delhi 110042, India; sumedhaseniaray@dtu.ac.in
3 Sustainable Manufacturing Systems Centre, School of Aerospace, Transport, and Manufacturing, Cranfield University, Bedfordshire MK43 0AL, UK
* Correspondence: pra.9000@gmail.com

Received: 2 August 2020; Accepted: 15 September 2020; Published: 17 September 2020



Abstract: Current frugal design practice is focused on the cost reduction of the product. Despite advancements in the domain of frugal Innovation, it is not systematized to develop products for all sets of users, including marginalized society. Many design researchers and engineers now dedicate time and knowledge to producing practical solutions to enhance the quality of life of the marginal community. The approach currently being adopted restricts the development of products intended for all segments of the users. In this paper, cumulative frequency distribution analysis and the Relative Importance Index is used to identify the essential attributes, which contribute to delivering actual frugal products in terms of functionality, usability, performance, affordability, accessibility, aesthetics, and robustness. The framework is beneficial to eradicate the discriminatory effect of being labeled as "Jugaad" users.

Keywords: frugal design; usability; functionality; reverse innovation

1. Introduction

Frugal design (FD) is one of the critical areas to fulfill the needs of the people in a resource-limited society. Frugal innovation (FI) has been identified as an essential element in improving the quality of life economically as well as socially. The current product development process in FD is primarily focused on (i) elimination of components, (ii) compromise with the quality (iii) reducing functionality, usability and aesthetic appeal, to reduce cost. There are a number of FI definitions available in the literature, and the most commonly cited one is the following: "Frugal innovation is a process of reducing the complexity and cost of a good and its production." Usually, this refers to removing nonessential features from a durable good to sell it in developing countries or low to mid-income market segments. The objective of FI is to reduce the complexity, overall cost, and enhance the functionality of the product, service, and systems while delivering high user value and affordable solutions within the economic range [1]. The process associated with mainstream or specialized products is directed towards specific audiences. FD processes focus on a unique set of users, i.e., people who are at the bottom of the pyramid, along with developing and undeveloped nations such as India, Nepal, Africa, etc. The outcome of the FD methodology would not be useful even if the product was aesthetically not attractive to all users, including mainstream users.

In this work, seven key elements (functionality, usability, performance, affordability, accessibility, aesthetics and robustness) for frugal innovation design are identified with the help of an cumulative frequency distribution analysis. A further descriptive-statistical tool called the Relative Importance Index (RII) is used, which helps to gain relative importance for each of these vital attributes identified

individually. RII is evaluated by obtaining the responses from two sets of respondents, design experts, and design practitioners, on a five-point Likert scale. Finally, the mean RII of every attribute is investigated based on the categorization of attributes, i.e., most essential, essential, desirable, and neutral contributing towards designing a frugal innovation. The main objective of this study is to draw an inference that the current FD practice is sufficient to develop frugal products as per the definition of FI and to introduce essential attributes to improve the FD practice to cater to all segment of users. A case study was conducted with the help of the Pugh chart method to gauge the frugality based on the identified attributes.

2. Background Study

Innovation is a popular subject of research in organizational philosophy [2]. It is identified as one of the prime indicators in obtaining a competitive advantage [3,4]. The nature and importance of innovation has been highlighted and discussed by [5–7]. In contrast to the developed economies, the developing economies implement a bottom-up strategy to the technology development by the reverse engineering of the commercial products. This gives rise to an unconventional method of product development where the information flows in a reverse direction (from distribution to the manufacturing to the research) [8]. This reverse trend is referred to as 'Frugal Innovation'. Such innovations are typically aimed at obtaining a low cost alternative to the existing technology and thus lack in meeting the intended functionality of the product [9,10]. There exists various versions of FD in the literature, namely, frugal engineering, reverse innovation, jugaad, constraint-based innovation [3,11,12]. However, a standardized conceptual approach is still missing in the literature. Immelt et al. [13] distinguished FI from the reverse innovation as the FIs are intended towards niche consumers and are not aimed for mass production. On the contrary, reverse innovation aims at developing profit-oriented businesses aiming at the worldwide distribution. This argument is further supported by Singhal [14], in that the frugal innovations are the innovations which are only restricted to their originating countries and are not intended for a global business.

Brem and Ivens [15] advocated the philosophy of the frugal innovation by stating that such designs involve simplicity at the same time imbibing functionality of the product. The authors further claim that FIs are sustainable in nature and utilizes less material and resources. The market performance of such innovations is directly linked with the degree of improvements achieved by the manufacturer. Weyruch and Herstatt [16] discussed three main characteristics of a frugal innovation: (a) significant decrease in the costs, (b) focus on product functionality, and (c) optimal performance. Figure 1 shows two examples of frugal innovation: Mitticool and VScan. Mitticool is an inexpensive refrigerator for Indian villages and towns with energy shortages. The cost of the refrigerator is 60% less compared to a commercial refrigerator. It can primarily be used for cooling water, dairy products, vegetables and fruits, thereby performing its intended function. However, the refrigerator lacks the freezer compartment and lightning. As it satisfies all the three criteria mentioned above, it can be considered as frugal innovation. Similarly, Vscan is a handheld scanner developed for the Chinese markets. The cost of this ultrasound scanner is 85% lesser than the existing commercial product. The core-functionalities of an ultrasound scanner are matched and the performance level also fits to the intended purpose. Thus, this can also be classified as an example of a FI.

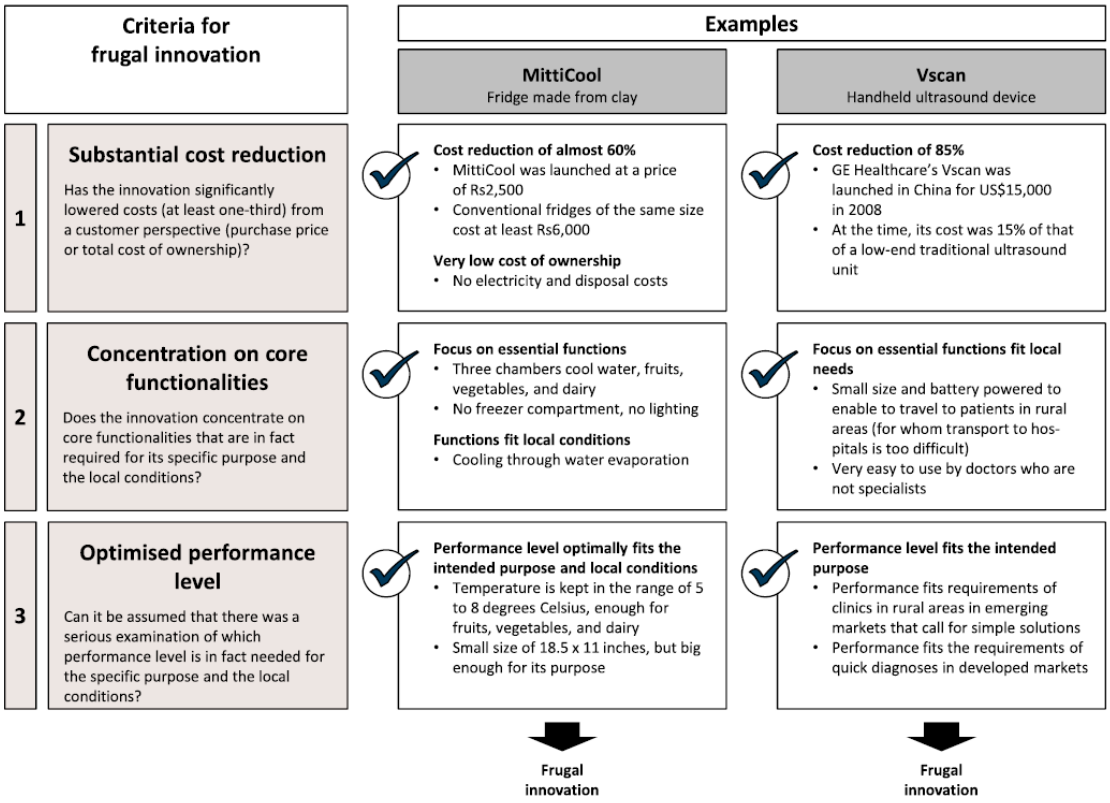


Figure 1. Examples of frugal innovation. Image adapted from [16].

Though FIs are mainly targeted at low-income consumers in developing countries, a few of these pass through to developed countries [17]. Some researchers claimed that FIs seek to challenge current product design frameworks with substantial cost savings on the one hand, while retaining customer satisfaction on the other [18]. The features of FIs are distinct from those of mainstream products, and they need to be viewed from a novel viewpoint [19]. Therefore, FIs require specific explanations for a variety of main reasons, such as:

1. Distinctive geographical background of the FIs [20]. Globalization is increasing the status of the middle class in developing economies, whose demand for better living conditions places our resources under pressure.
2. Unique distribution method [21]. For example, FIs can be distributed and sold by the local community using the FIs, who are also trained as salespersons. This can help in bringing down the distribution costs.
3. They ought to have a new business model [9]. Traditionally, business models of firms in the emerging countries do not focus on the resource-constrained customers, but rather are concentrated on the privileged few who have the higher purchasing power. Western businesses wishing to participate in frugal innovation need to create operational frameworks and skills to produce frugal goods.
4. Unique distribution channel is needed [22].
5. Crude design of the FIs compared to the mainstream products, as the FIs are more focused on the low cost of the intended product.
6. Difference in aesthetics and ergonomics of the FIs.

3. Method

To develop frugal products, designers need to acknowledge users' needs and wants during the product development phase. The ultimate goal of this study is to determine the essential attributes of

FI that can be incorporated during the early phase of design process. The method of determining the attributes consist of three steps (Figure 2), as given below:

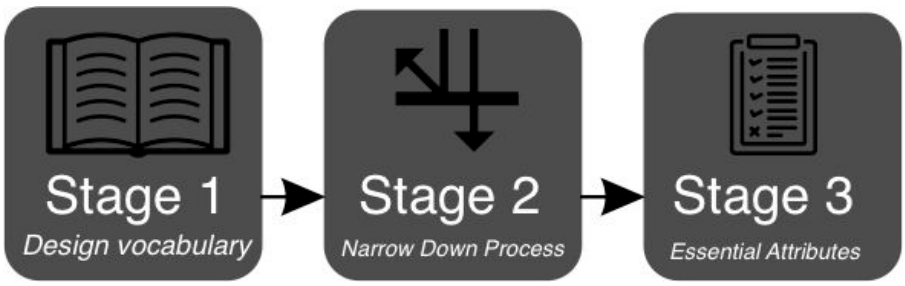


Figure 2. Methodology.

3.1. Design Vocabulary

A list of attributes that are associated with the design and users were collected from the journals, books, internet, design magazines, product catalogues and market survey. A list of more than a hundred words were identified and collected for the study, representing the design attributes based on the tangible/in tangle outcomes and user characteristics (e.g., functionality and usability extracted from tangible/in tangle outcomes i.e., product and product attachment/emotion is obtained from human experience). The objective of this step was to explore and create design vocabulary to identify the required attributes (Appendix A) [23–25].

3.2. Narrow Down Processes

The second step was to condense the database by eliminating the inappropriate words that were not directly associated with FI. All words that did not fall into the category of the product feature or attribute were excluded from the database (e.g., charity, health, heart, devotion, energy, etc.). Finally, the elimination of similar words was done. The number of attributes was condensed to more than twenty by neglecting words that appeared improper for expressing frugality. In this process, 30 design experts from academia and industry were involved to identify the relevant attributes by removing similar and inappropriate attributes from the design vocabulary.

The total number of attributes was narrowed down to 11 attributes, as shown in Table 1, after the focus group session with the design experts in which words are mapped with the definition of frugal design. Cumulative frequency distribution analysis was performed to identify the important attributes that are directly or indirectly related to frugality with the help of 30 design experts (Table 2). Cumulative frequency is an important tool in statistics, and also in data analytics; it helps to govern the number of observations that lie above (or below) a certain value in our dataset [26]. It is also used to extract the information from the dataset, i.e., decision making, trend analysis, forecasting, etc. Cumulative frequency distribution was tabulated with the help of distinct attributes and their corresponding occurrences. The analysis table based on cumulative frequencies was presented by the steps stated:

1. Absolute frequencies for each of the attributes were collected from the design experts.
2. Frequencies were arranged in descending order.
3. The cumulative frequency for each corresponding design attribute was further calculated.
4. Finally, the percentage corresponding to each attribute was calculated using the following formula:

$$Percentage\ (\%) = \frac{c\,f}{\sum_{i=1}^{n} f_i} * 100$$

where

cf = Cumulative frequency

$\sum f_i$ = Sum of absolute frequencies
n = Total number of attributes

Table 1. Attributes associated to frugality.

Filtered Attributes		
Functionality	Usability	Affordability
Accessibility	Performance	Aesthetics
Robustness	Human factors	Adaptability
Modularity	Attachment	

Table 2. Cumulative frequency distribution for predicting vital attributes.

S.No.	Attributes	Frequency	Cumulative Frequency	Percentage
1	Functionality	30	30	16.13
2	Usability	23	53	28.49
3	Affordability	23	76	40.86
4	Accessibility	22	98	52.69
5	Performance	21	119	63.98
6	Aesthetics	17	136	73.12
7	Robustness	13	149	80.11
8	Human factors	11	160	86.02
9	Adaptability	10	170	91.40
10	Modularity	10	180	96.77
11	Attachment	6	186	100.00

The cumulative frequency distribution analysis was applied to eliminate the uncertainties related to design attributes and to prioritize the root-cause and/or problem-solving. The cumulative frequency distribution chart, therefore, shows the relative frequency design attributes in a rank-order, and thus delivers a prioritization tool so that process improvement activities can be organized to get the most effective attributes towards the frugal innovation from others. The purpose of applying the cumulative frequency distribution analysis was to separate the "vital attributes" from the "trivial many", which was represented with the help of the cumulative frequency distribution chart.

Figure 3 and Table 2 shows the absolute frequency and cumulative frequency distribution of the 11 design attributes. After evaluating the cumulative frequencies and their corresponding percentage, we applied an evaluation rule on these attributes, which states that:

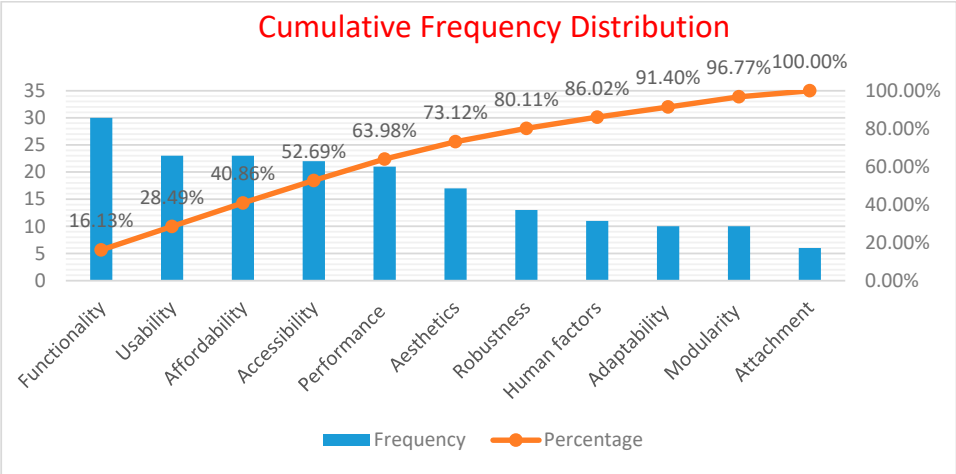


Figure 3. Cumulative frequency distribution analysis chart.

"All the attributes lying within 80% are recognized to be the vital attributes contributing towards frugal innovation".

This evaluation rule acts as a threshold in the cumulative frequency distribution analysis, predicting which among all the attributes lies above the stated threshold value, are recognized as the significant ones. Therefore, based on this evaluation rule, 7 attributes were recognized to significantly contributing towards frugal innovation, namely:

1. Functionality: The quality of being suited to serve a purpose well [27].
2. Affordability: Affordability is related to an economical bond that consumer experiences during the purchase of the product at an efficient price.
3. Usability: Ease of use and learnability of human made products.
4. Aesthetics: Concerned with emotions which evoke a sense of beauty and precision in relation to form, color, material and finish.
5. Robust: Product, service, and environment able to withstand or overcome adverse conditions and perform effectively.
6. Performance: The action or process of performing a task or function [27].
7. Accessibility: Products and environments should be designed to be usable by all sets of users without modifications.

Thus, the seven important attributes were pinned down for the development of frugal designs.

3.3. Essential Attributes

To determine the importance of each of these seven vital attributes towards a frugal innovation, a descriptive-analytical tool called the Relative Importance Index (RII) on the basis of a Likert scale was employed. The Relative Response Index is used for performing regressive analysis on the attributes that have been identified in the previous stage of cumulative frequency distribution. It acts as a predictor, which helps to gain insight to which of the given factors are the most important or are of lesser importance. In addition, with the help of RII, we can compare the responses from each of the group of respondents in order to draw a comparison among the factors, indicating which factor is more important to either of the sets of respondents in an ordered custom. In order to draw this conclusion, we again deployed a panel of design experts and design practitioners who are actively associated with the process of design innovation, that helped us determine the importance of these attributes, which was achieved by collecting individual responses on a 5-point Likert scale. There were three steps in carrying out this analysis, which are discussed in the subsequent sections.

3.3.1. Step I: Data Collection through Survey

In this step, a survey was conducted with 15 design experts and 25 design practitioners where they had to rate the seven attributes on the concept of importance for each of these attributes on a 5-point Likert scale of 1 to 5, 1 being the least important attribute and 5 being the most important attribute. Tables 3 and 4 represents the expert and end-user responses.

Table 3. Factors affecting designing of a frugal innovation (design expert responses).

	Response Per Frequency					
Attributes	**5—Most Important**	**4—Important**	**3—Moderately Important**	**2—Least Important**	**1—Not Important**	**Total Respondent**
Functionality	11	4	0	0	0	15 Design Experts
Usability	4	8	3	0	0	
Affordability	13	1	1	0	0	
Accessibility	0	2	7	5	1	
Performance	0	8	6	1	0	
Aesthetics	5	5	4	0	1	
Robustness	5	6	3	1	0	

Table 4. Factors affecting designing of a frugal innovation (design practitioner responses).

	Response Per Frequency					
Attributes	**5—Most Important**	**4—Important**	**3—Moderately Important**	**2—Least Important**	**1—Not Important**	**Total Respondent**
Functionality	14	10	1	0	0	25 Design practitioners
Usability	10	11	3	1	0	
Affordability	10	6	9	0	0	
Accessibility	2	6	9	6	2	
Performance	8	9	8	0	0	
Aesthetics	11	9	5	0	0	
Robustness	4	10	9	2	0	

3.3.2. Step II: Calculating Relative Importance Index (RII)

The data received by both set of respondents were then analyzed with the help of a descriptive statistical tool RII in order to determine the relative importance of each of these factors or attributes corresponding to a frugal innovation.

Mathematically, the Relative Importance Index is evaluated as follows:

$$\text{RII} = \frac{\sum_{i=1}^{5} W_i n_i}{A \times N}$$

where

RII = Relative Importance Index
W_i = Weights provided by ith respondent to each factor, ranging from i = 1 to 5
n_i = Total number of responses in each response category of the Likert scale
A = Highest weight in the Likert scale (i.e., "5")
N = Total number of respondents

Based on the responses from the design experts and the design practitioners, the RII evaluation for each of the factors from both sets of respondents is represented in Table 5.

Table 5. Evaluated Relative Importance Index (RII) values of the respondents.

Attributes	**Expert RII**	**Design Practitioner RII**
Functionality	0.947	0.904
Usability	0.960	0.808
Affordability	0.813	0.840
Accessibility	0.773	0.848
Performance	0.800	0.728
Aesthetics	0.693	0.800
Robust	0.533	0.600

3.3.3. Step III: Calculating Mean RII

To determine the contribution of each of these factors towards frugal innovation based on its importance, mean RII was calculated [28] from the RII values obtained from Table 5. Figure 4 illustrates the mean RII for each of these factors, and they are further categorized as most essential, essential, desirable, and neutral (refer Table 6 andTable 7).

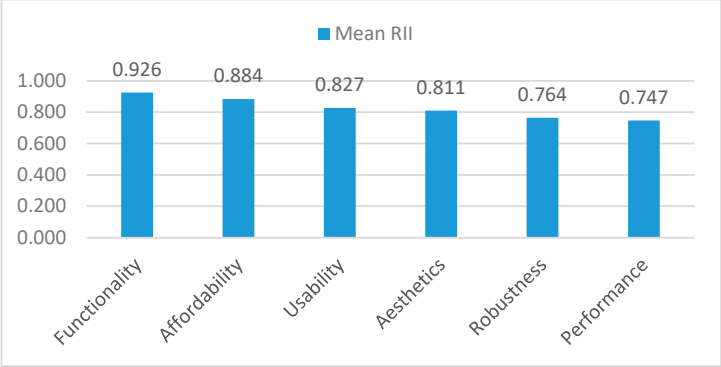


Figure 4. Mean Relative Importance Index (RII).

Table 6. Evaluation rule.

Attributes	Mean RII Range
Most essential	$(0.8 < RII \leq 1.0)$
Essential	$(0.6 < RII \leq 0.8)$
Desirable	$(0.4 < RII \leq 0.6)$
Neutral	$(RII \leq 0.4)$

Table 7. Mean RII values and corresponding categories.

Attributes	Mean RII	Category
Functionality	0.926	Most Essential
Affordability	0.884	Most Essential
Usability	0.827	Most Essential
Aesthetics	0.811	Most Essential
Robustness	0.764	Essential
Performance	0.747	Essential
Accessibility	0.567	Desirable

The importance of each of these factors was observed and recorded in Table 7, concluding that functionality, affordability, usability, and aesthetics are the most essential attributes. Robustness and performance represented essential attributes of frugal innovation. Attributes such as accessibility, falling under desirable attributes, will enhance the overall experience within the diverse population.

4. A Case Study of the Passenger Cars Based on Attributes

A case study of the passenger cars was then carried out to validate the method. Two economical passenger cars (Tata Nano and Renault Kwid—shown in Figure 5) were taken for this case study. These cars were chosen to investigate the product, based on the essential parameters of the FIs as the selected product served well for users belonging to the low and the middle income groups.



(**a**) (**b**)

Figure 5. Passenger car case study: (**a**) Tata Nano; (**b**) Renault Kwid.

The attributes associated to the frugal innovation such as explanation or details of the attributes were provided to the participants. The Likert-type questionnaire was handed over to the users and their response was recorded based on their experience and perception of using the selected product on a scale of 5, with the following labels: '++'excellent; '+' good; '0' average; '−' poor; and '−−'worse.

After investigating the responses in Table 8, it was observed that the TATA Nano performed well in terms of affordability, whereas the Renault Kwid performed extremely well in terms of aesthetics, usability, functionality, and affordability.

Table 8. Case study (i) TATA Nano, and (ii) Renault Kwid.

		Attributes	**Tata Nano**	**Renault Kwid**
7 Attributes of Frugal Innovation	**5 Expert Users**	Functionality	+	++
			++	++
			++	+
			0	+
			+	++
		Usability	+	++
			+	++
			+	++
			+	+
			0	++
		Affordability	++	++
			++	+
			++	++
			++	++
			++	+
		Accessibility	0	0
			0	0
			0	+
			0	−
			0	0
			−	
		Performance	0	0
			0	+
			+	+
			−	+
			+	0
		Aesthetics	+	++
			++	++
			++	+
			0	++
			−	++
		Robust	+	+
			+	+
			+	0
			0	+
			0	+
Total "+"			30	41
Total "−"			3	1
Total Score			27	40

The case study of the passenger car was useful to examine what attributes led to the failure of the design in terms of frugality. Affordability is not the only criteria which defines the frugality—other attributes such as functionality, usability, performance, aesthetics, robust, accessibility are also essential attributes of frugal innovation. The current frugal design approach is more focused on cost reduction, functionality, and performance sustainability [29]. Although these are very important attributes to develop effective products in terms of economy, productivity, and ecology aspect of the design spectrum [30,31], social inclusion is completely neglected in current practice, and the inclusion of all sets of users into consideration will help to map the product on the social, ecological and economic scale, and enhance the acceptability among the diverse user sets. The product would be productive if users found its appearance to be elegant and comfy to use. Designers need to consider aesthetics during the initial design process as aesthetics plays a vital role to enhance the reachability and scalability of the product.

5. Conclusions

Frugal innovation is a professional design process that helps to reduce the complexity and cost of the product while at the same time increasing the efficiency of the product. The identification of essential attributes is the initial step towards true frugal innovation. This study identifies the seven essential attributes of frugal innovation with the help of 80/20 analysis and the Relative Importance Index (RII). A case study was performed to investigate the attributes of FI that helped us to analyze whether the identified attributes from our experiment are true to their cause for designing a frugal innovation. TATA Nano was the case study that was undertaken which turned out to be an affordable product but not effective in terms of aesthetics, usability, functionality, and affordability. Negligence of these attributes during the initial design phase creates the prejudicial effect of being labeled as a low-cost product "Jugaad", which conflicts with the objective of frugality. The product would be productive if users found its appearance to be elegant and comfy to use. There are numerous products under the label of "Jugaad" but are not considered mainstream products among the users; because of a perception of disgrace/awkwardness attached to them. Therefore, in many instances, even marginalized communities avoid using the product that is specially designed for them. Designers need to consider aesthetics during the initial design process, as aesthetics plays a vital role to enhance the reachability and scalability of the product. Thus, based on the observation and the case study carried out, it can be concluded that if the designer considers these seven attributes in the initial design phase of the product design process, the product would be frugal in every sense—that is, in terms of cost, function and aesthetics. The proposed framework is effective to eradicate the discriminatory effect of being labeled as "Jugaad" users.

Author Contributions: Conceptualization, R.S. and P.S.; methodology, R.S.; background study, P.S.; case study and validation, S.S.; writing, review and editing, all. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. List of attributes.

Human factors, Innovation	Refinement	Economy
Usefulness	Recreation	Consistency
Usability	Relaxation	Capability
Aesthetics	Desire	Bliss
Self-explanatory	Energy	Justice
Affordance	Respect	Maturity
Honest	Sympathy	Control
Simple	Relief	Contribution
Comfort,	Hope	Capability
Acceptance,	Want	Energy
Redundancy	Adaptability	Security
Balance	Challenge	Self-reliance
Accessibility	Fitness	Stability
Capacity	Flexibility	Support
Accuracy	Cheerfulness	Surprise
Intuitive	Friendliness	Amused
Flexibility	Frugality	Peaceful
Feedback	Fun	Alert
Adaptability	Direction	Attachment
Modularity	Commitment	Perfect
Interactive	Confidence	Helpfulness
Performance	Appreciation	Honor
Functionality	Availability	Hopefulness
Product attachment	Awareness	Courage
Intuition	Beauty	Courtesy
Elegance	Connection	Order
Empathy	Gratitude	Originality
Enjoyment	Growth	Motivation
Entertainment	Guidance	Courage
Hospitality	Vision	Uniqueness
Creativity	Unity	Proactive
Hygiene	Prosperity	Richness
Imagination	Professionalism	Utility
Impact	Resolve	
Excitement	Reliability	
Pleasure	Simplicity	
Insightfulness	Satisfaction	
Inspiration	Togetherness	
Understanding	Independent	
	Aspiration	

References

1. Leliveld, A.; Knorringa, P. Frugal Innovation and Development Research. *Eur. J. Dev. Res.* **2017**, *30*, 1–16. [CrossRef]
2. Brem, A. Linking innovation and entrepreneurship—Literature overview and introduction of a process-oriented framework. *Int. J. Entrep. Innov. Manag.* **2011**, *14*, 6–35. [CrossRef]
3. Drazin, R.; Schoonhoven, C.B. Community, Population, and Organization Effects on Innovation: A Multilevel Perspective. *Acad. Manag. J.* **1996**, *39*, 1065–1083.
4. Christensen, C.M. *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*; Harvard Business Review Press: Boston, MA, USA, 2013.
5. Chandy, R.K.; Hopstaken, B.; Narasimhan, O.; Prabhu, J.C. From Invention to Innovation: Conversion Ability in Product Development. *J. Mark. Res.* **2006**, *43*, 494–508. [CrossRef]

6. Jon-Arild, J.; Bjørn, O.; Lumpkin, G.T. Innovation as newness: What is new, how new, and new to whom? *Eur. J. Innov. Manag.* **2001**, *4*, 20–31.
7. Govindarajan, V.; Trimble, C. Organizational DNA for strategic innovation. *Calif. Manag. Rev.* **2005**, *47*, 47–76. [CrossRef]
8. Gardner, P.L. The Globalization of R&D and International Technology transfer in the 21st century. In Proceedings of the International Conference of Management of Innovation and Technology (ICMIT'02 & ISMOT'02), Hangzhou, China, 18–20 October 2002; Volume 3, p. 5.
9. Zeschky, M.; Widenmayer, B.; Gassmann, O. Frugal Innovation in Emerging Markets. *Res. Manag.* **2011**, *54*, 38–45. [CrossRef]
10. Agarwal, N.; Brem, A. Frugal and reverse innovation-Literature overview and case study insights from a German MNC in India and China. In Proceedings of the 2012 18th International ICE Conference on Engineering, Technology and Innovation, Munich, Germany, 18–20 June 2012; pp. 1–11.
11. Saraf, D. India's Indigenous Genius: Jugaad—WSJ. 2009. Available online: https://www.wsj.com/articles/SB124745880685131765 (accessed on 29 July 2020).
12. Singh, R.; Das, P.P. An Approach to Develop Accessible and Affordable Products. *Procedia CIRP* **2020**, *91*, 489–493. [CrossRef]
13. Immelt, J.R.; Govindarajan, V.; Trimble, C. How GE is disrupting itself. *Harv. Bus. Rev.* **2009**, *87*, 56–65.
14. Singhal, V. The impact of emerging economies innovative new models of global growth and vitality are emerging. *Visions* **2011**, *35*, 12–14.
15. Brem, A.; Ivens, B.S. Do Frugal and Reverse Innovation Foster Sustainability? Introduction of a Conceptual Framework. *J. Technol. Manag. Grow. Econ.* **2013**, *4*, 31–50. [CrossRef]
16. Weyrauch, T.; Herstatt, C. What is frugal innovation? Three defining criteria. *J. Frugal Innov.* **2016**, *2*, 1. [CrossRef]
17. Von Zedtwitz, M.; Corsi, S.; Søberg, P.V.; Frega, R. A typology of reverse innovation. *J. Prod. Innov. Manag.* **2015**, *32*, 12–28. [CrossRef]
18. Knorringa, P.; Peša, I.; Leliveld, A.; Van Beers, C.; Knorringa, I.P.P. Frugal innovation and development: Aides or adversaries? *Eur. J. Dev. Res.* **2016**, *28*, 143–153. [CrossRef]
19. Hossain, M. Frugal innovation: A review and research agenda. *J. Clean. Prod.* **2018**, *182*, 926–936. [CrossRef]
20. Rao, B.C. How disruptive is frugal? *Technol. Soc.* **2013**, *35*, 65–73. [CrossRef]
21. Hossain, M.; Simula, H.; Halme, M. Can frugal go global? Diffusion patterns of frugal innovations. *Technol. Soc.* **2016**, *46*, 132–139. [CrossRef]
22. Simula, H.; Hossain, M.; Halme, M. Frugal and reverse innovations quo vadis? *Curr. Sci.* **2015**, 1567–1572. [CrossRef]
23. Falvo, D.A.; Urban, M.; Lidwell, W.; Holden, K.; Butler, J. Universal Principles of Design. *Educ. Technol. Res. Dev.* **2007**, *55*, 297–300. [CrossRef]
24. Desmet, P.M.A. Faces of product pleasure: 25 positive emotions in human-product interactions. *Int. J. Des.* **2012**, *6*, 2.
25. Desmet, P.; Hekkert, P. Framework of product experience. *Int. J. Des.* **2007**, *1*, 57–66.
26. Chakraborty, P. Histogram equalization by cumulative frequency distribution. *Int. J. Sci. Res. Publ.* **2012**, *2*, 7.
27. Singh, R.; Tandon, P. Framework for improving universal design practice. *Int. J. Prod. Dev.* **2018**, *22*, 377–407. [CrossRef]
28. Aswathi, P.; Wilson, A. Study on Critical Performance Factors Affecting Kochi Metro Rail Project. *Int. J. Adv. Sci. Res. Eng.* **2020**, *6*, 107–112.
29. Winkler, T.; Ulz, A.; Knöbl, W.; Lercher, H. Frugal innovation in developed markets—Adaption of a criteria-based evaluation model. *J. Innov. Knowl.* **2019**. [CrossRef]
30. Singh, R.; Tandon, P. A Sustainable Approach to Develop Universal Products. In Proceedings of the 6th International Kansei Engineering and Emotion Research Conference, Leeds, UK, 31 August–2 September 2016.
31. Saxena, P.; Stavropoulos, P.; Kechagias, J.; Salonitis, K. Sustainability Assessment for Manufacturing Operations. *Energies* **2020**, *13*, 2730. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

View publication stats

Neural Processing Letters
https://doi.org/10.1007/s11063-020-10345-1



A New Robust Fuzzy Clustering Approach: DBKIFCM

Anjana Gosain[1] · **Sonika Dahiya**[2]

Accepted: 1 September 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract
A clustering algorithm—Distance Based Gaussian Kernelized Intuitionistic Fuzzy C Means (DBKIFCM) is proposed. This algorithm is based on Gaussian kernel, outlier identification, and intuitionist fuzzy sets. It is intended to resolve the issue of presence of outliers, problem of sensitivity to initialization (STI) and is motivated by good performance of Radial Based Kernelized Intuitionistic Fuzzy C Means (KIFCM-RBF). Experiments are performed on standard 2D data sets such as Diamond (D12 and D15), and Dunn and real-world high dimension data sets such as Fisheriris, Wisconsin breast cancer, and Wine. DBKIFCM outcomes are studied in relation to Fuzzy C Means (FCM), Intuitionistic Fuzzy C Means (IFCM), KIFCM-RBF, Density Oriented Fuzzy C Means (DOFCM). It is observed that proposed approach significantly outperforms the earlier proposed algorithms with respect to outlier identification, effect of noise, issue of STI, and clustering error.

Keywords Fuzzy clustering · Outlier identification · Kernel function · FCM · IFCM · KIFCM · DOFCM

1 Introduction

Clustering is useful in identifying natural boundaries in data set whereas fuzzy clustering is useful in cases where cluster boundaries are not sharp. Fuzzy clustering deals with fuzziness, uncertainty and vagueness in data. Thus, it is extensively used in domains such as medical imaging [1], astronomy, finance, marketing, robust designing etc. [2]. Fuzzy clustering is applied in various applications such as object recognition, customer segmentation, fault diagnosis [3], pattern recognition, image segmentation [4–6].

In 1965, Lotfi Zadeh introduced fuzzy sets [7]. Dunn, in 1974, proposed ISODATA algorithm [8] by incorporating fuzzy logic in clustering and in 1981, an extension of ISODATA algorithm, FCM (Fuzzy C-Means) [9] was proposed by J.C. Bezdek which is widely accepted and successfully applied in fuzzy clustering applications. However, the performance of FCM is affected by the presence of noise and outliers. To smoothen the effect of noise, Tolia

✉ Sonika Dahiya
sonika.dahiya11@gmail.com

1 USICT, GGSIP University, Sector 16C, Dwarka, Delhi 110075, India

2 CSE, Delhi Technological University, Main Bawana Road, Shahbad Daulatpur, Delhi 110042, India

Published online: 20 September 2020

and Panas [10] suggested post-processing of membership function, introduced in FCM. For enforcing spatial constraint, Acton and Mukherjee [11, 12] incorporated multiscale information. In 1993, Krishnapuram et al. proposed Possibilistic C Means (PCM) [12, 13] which interprets clustering as a possibilistic partition. But PCM is sensitive to initialization and few data objects are allocated to more than one cluster, which causes cluster identity issues or overlapping issues. In 2005, Possibilistic Fuzzy C Means [12, 14] was proposed by Nikhil R. Pal et al. to overcome the identical cluster problem of PCM by generating typicality and membership values. However, it doesn't perform well for noisy and unequal sized clusters in data. Chintalapudi and Kam in 1998, proposed credibilistic fuzzy c-means (CFCM) [12, 15] by introducing credibility variable to deal with the presence of outliers. Though it is effective in reducing the effect of outliers on computing cluster centroids, but few outliers are associated with more than one cluster [8]. In 2011, P. Kaur et al. proposed DOFCM (density based FCM) [12, 16], which uses the concept of density and neighborhood membership to uniquely identify outliers.

To maximize the good points in a cluster, T.Chaira, in 2011, proposed Intuitionistic fuzzy C Means (IFCM) [12, 17] by incorporating hesitation degree with membership function and intuitionistic fuzzy entropy in objective function. It was successful in converging the cluster center to a desired location. In 2013, P. Kaur et al. proposed Kernelized Intuitionistic Fuzzy C-Means (KIFCM-RBF) [1], an extension of IFCM [18], by incorporating kernel function for distance measurement in high dimension feature space. This improves clustering results especially for nonlinearly separable clusters. However, algorithm efficiency degrades significantly in the presence of outliers.

In this paper, we propose a Distance Based Gaussian Kernel IFCM (DBKIFCM) algorithm which uses a distance-based approach as a preprocessing step for outlier identification and centroid initialization. In preprocessing step, we compute outlier spatium (spatium: Latin word which means distance) and uses this outlier spatium for identification of outliers. DBKIFCM initializes centroids to those data objects that have highest count in spatium neighborhood and uses Gaussian kernel function to map data to high dimension feature space and then performs clustering using intuitionistic fuzzy c-means. Experiments are performed on synthetic 2D data sets such as Diamond data set (D12 and D15), Dunn data set and on high dimension real data sets such as Fisheriris, Wisconsin Breast Cancer, Wine. It is found that DBKIFCM outperforms earlier proposed clustering algorithms by improving clustering results, centroid position for linear data sets, non-linear data sets and data sets with outliers.

This paper is organized as: Brief review of FCM, IFCM, KIFCM(RBF), DOFCM in section II; Description of the proposed algorithm DBKIFCM in section III; Experimental evaluation and results on standard 2-dimensional and higher dimensions data sets in section IV; and Conclusion in section V.

2 Related Work

In this section, we briefly discuss FCM, IFCM, KIFCM and DOFCM. Following notations are used: $D = \{d_1, d_2, d_3, \ldots, d_n\}$ represents set of data objects where d_j is j^{th} data object; C_i is i^{th} centroid; md_{ji} is fuzzy membership degree of d_j in C_i; $dist_{ji}$ is Euclidean distance between d_j and C_i; nc is count of clusters; n is count of data objects; m is fuzziness index [8, 9].

2.1 The FCM [9]

It is one of the widely accepted fuzzy clustering algorithm [9]. It works on assumption that count of clusters is known and minimizes the following objective function (J_{FCM}):

$$J_{FCM} = \sum_{i=1}^{nc} \sum_{j=1}^{n} md_{ji}^{m} dist_{ji}^{2} \tag{1}$$

Subject to following constraint:

$$\sum_{i}^{nc} md_{ji} = 1 \quad \forall j \in [1, n] \tag{2}$$

For our simulation m is set to be 2. On solving optimization problem stated in Eq. (1), following equations are found to be used for membership and centroid updation:

$$md_{ji} = \frac{1}{\sum_{k=1}^{nc} \left(\frac{dist_{ij}}{dist_{kj}} \right)^{\frac{2}{(m-1)}}} \quad \forall j \in [1, n] \tag{3}$$

$$C_i = \frac{\sum_{j=1}^{nc} \left(md_{ji}^{m} d_j \right)}{\sum_{i=1}^{nc} md_{ji}^{m}} \tag{4}$$

where i and j are positive integers in [1, nc] and [1, n] respectively.

However, it proves to be insufficient in the presence of noise and outliers, as it fails to identify noise and outliers. So, the centroids are inclined towards outliers. Even with these issues, FCM has wide applications such as medical image segmentation, defect identification etc. [19–22].

2.2 Intuitionistic Fuzzy C Means (IFCM) [18, 23]

Xu and Wu based on intuitionistic fuzzy set theory [8] proposed IFCM [18] which is helpful in dealing with uncertain and vague data [18, 23]. Its objective function is:

$$J_{IFCM} = \sum_{i=1}^{nc} \sum_{j=1}^{n} md_{ji}^{*m} \cdot dist_{ji}^{2} + \sum_{j=1}^{nc} h_j^{*} e^{1-h_j^{*}} \tag{5}$$

with m is set to be 2, $md_{ji}^{*} = md_{ji} + \eta_{ji}$, where md_{ji}^{*} represents intuitionistic fuzzy membership [18, 23] and η_{ji} is hesitation degree [23] defined as:

$$\eta_{ji} = 1 - md_{ji} - \left(1 - md_{ji}^{\alpha}\right)^{\frac{1}{\alpha}} \tag{6}$$

And

$$h_j^{*} = \frac{1}{n} \sum_{i=1}^{n} \eta_{ji} \tag{7}$$

IFCM produces overlapping clusters and hence it becomes really difficult to assign a cluster to the points lying in overlapping region. Also, IFCM fails to handle outliers as this algorithm treats outliers as data objects.

2.3 Kernel Version of Intuitionistic Fuzzy C Means (KIFCM) [24]

KIFCM make use of Radial Basis kernel function [25] for computing the distance between the data objects and centroid. Objective function is:

$$J_{KFCM} = \sum_{i=1}^{nc}\sum_{j=1}^{n} md_{ij}^{*m} \left\| \Phi\left(d_j\right) - \Phi(C_i) \right\|^2 + \sum_{i=1}^{nc} h_i^* e^{1-h_i^*} \tag{8}$$

2.4 Density Oriented Fuzzy C Means (DOFCM) [26, 27]

In 2010, Kaur et al. proposed DOFCM by introducing a new term neighborhood membership, defined as follows:

$$M^i_{neighborhood}(X) = \frac{\eta^i_{neighborhood}}{\eta_{\max}} \tag{9}$$

where $\eta^i_{neighborhood}$ = count of data objects in d_i neighborhood and $\eta_{\max} = \underset{i=1..n}{\max}\left(\eta^i_{neighborhood}\right)$.Any data object '$d_p$' is in neighborhood of d_i if it satisfies following condition:

$$p \in X,\ i \in X \left| dist(p,i) \le r_{neighborhood} \right. \tag{10}$$

where $r_{neighborhood}$ is neighborhood radius.

To distinguish outliers from normal data objects, a threshold value of neighborhood membership is defined as 'α' and data objects(di) whose $M^i_{neighborhood}$ value falls below threshold are identified as outliers. Post this, FCM clustering is performed. DOFCM is only fuzzy clustering approach that focuses on identification of outliers. Thus, it has many real life applications such as fault diagnosis [28], online fault detection [29, 30].

3 Proposed Work

3.1 Distance Based Gaussian Kernelized Intuitionistic Fuzzy C Means (DBKIFCM)

After explaining IFCM, KIFCM(RBF) and DOFCM, now we are able to construct a distance-based gaussian kernel version of intuitionistic fuzzy clustering named as DBKIFCM. The proposed algorithm first identifies outliers and highlights them explicitly by assigning zero membership. It then performs clustering in high dimension feature space in such a way that nullifies presence of outliers and results in more accurate and noise free clusters. In the following sections we explain how outliers are identified, how kernelization approach works and how proposed clustering algorithm works.

3.2 Outlier Identification Approach

An outlier is basically any kind of anomaly in the data with respect to behavior or context of data objects, making outlier study very significant. Outlier Identification in itself is a major research area in Data Mining as presence of outliers significantly degrades performance of clustering. In literature DOFCM [8, 12] has been the best fuzzy clustering algorithm which

works by effectively identifying the outliers and then performs clustering. But it has following drawbacks: (1) focuses only on local outliers, (2) quite sensitive to the choice of threshold value, (3) does not converge well for threshold value.

We have proposed a new outlier identification technique that have overcome all the drawbacks of DOFCM. Outlier identification of DBKIFCM is a distance-based approach. For a given data set, a data object is identified as outlier if an assumed fraction of data (f_D) is not present within a well computed outlier spatium ($r_{Outlier_Spatium}$). Mathematical formulation of outlier identification is as follows:

$$knn = \lceil \|D\|.f_D \rceil \tag{11}$$

$$\left(\left\|d_j \middle| dist(d_i, d_j) \leq r_{Outlier_Spatium}\right\|\right) < \text{knn} \quad \text{for } \{\forall j \in [1..N],\ j \neq i\} \tag{12}$$

where knn is k-nearest neighbors computed as per Eq. (11), f_D is a fraction of data set, and ||D|| is size of data. In Eq. (12), d_i and d_j are ith and jth data object of the data set D and $dist(d_i, d_j) = \sqrt{\sum (d_{iz} - d_{jz})^2}$ represents Euclidean distance between data object d_i and d_j as discussed in Sect. 2. $r_{Outlier_Spatium}$ is outlier spatium, which is computed using above defined knn equation and Ester et al. [31] defined heuristic approach. As per Ester et al. [31], if dt is the distance between a 1

data object 'pt' and its knn th nearest neighbor, then circumscribed neighborhood of 'pt' with radius 'dt' would have exactly knn + 1 points in its perimeter. The neighborhood for any data object, pt, would contain more than knn + 1 data objects, only if many data objects have exactly the same distance dt from pt. However, this scenario is highly unlikely. Also, changing knn for any data object in a cluster does not results in large variation of dt, except the case in which all knnth data points are located approximately on a straight line. However, such straight line case is highly unlikely in clustering. So, based on above heuristic approach, Ester [31] gave a method to compute dt, keeping value for knn as 4 based on their experiments. Similarly, we computed $r_{Outlier_Spatium}$ using knn as per Eq. (11). From Eqs. (11) and (12):

$$\left(\left\|d_j \middle| dist(d_i, d_j) \leq r_{Outlier_Spatium}\right\|\right) < \lceil \|D\|.f_D \rceil \text{ for } \{\forall j \in [1..N],\ j \neq i\} \tag{13}$$

Any data object 'd_j' is an outlier if it fails to hold at least as many data objects as ||D|| multiplied by fraction of data (f_D) in its proximity area. Proximity area is defined as space drawn taking d_j as center and $r_{Outlier_Spatium}$ as radius. For a given data set, complete range of outlier spatium ($r_{Outlier_Spatium}$) depends on f_D. A threshold value of f_D is selected depending on visual analysis of data set.

$$d_j = \begin{cases} outlier \text{ if } \left(\left\|d_j \middle| dist(d_i, d_j) \leq r_{Outlier_Spatium}\right\|\right) < \text{knn} \quad \text{for } \{\forall j \in [1..N],\ j \neq i\} \\ not \text{ outlier if } \left(\left\|d_j \middle| dist(d_i, d_j) \leq r_{Outlier_Spatium}\right\|\right) \geq \text{knn} \quad \text{for } \{\forall j \in [1..N],\ j \neq i\} \end{cases} \tag{14}$$

3.2.1 Selection of f_D

An ideal outlier is a data object which is in complete isolation to all other data objects i.e. knn for it would be zero. However, in our proposed approach, a data object is said to be an outlier only if it fails to hold at least a certain fraction of data in its proximity area. Here, proximity area is a very critical parameter for outlier identification. Its value depends on assumed f_D, spread, size, and features of data set such as density of a data set. So, its value varies for different data sets.

Table 1 Effect of various choices of f_D

# Data objects in proximity area	f_D	#Outliers	#Data objects	knn	$r_{Outlier_Spatium}$	Sub plot in Fig. 1
1	0.0222	13	32	1	4.2168	(a)
2	0.0444	14	31	2	5.9634	(b)
3	0.0667	12	33	3	8.4335	(c)
18	0.4000	12	33	18	17.8903	(d)
19	0.4222	19	26	19	18.3805	(e)
20	0.4444	30	15	20	18.8500	(f)

Explanation with Example Figure 1 and Table 1 illustrates working logic of proposed approach on D45, where D45 is a data set containing 45 data objects. Outliers are represented by red dot and proximity area is drawn as circle with outlier as center and radius as $r_{Outlier_Spatium}$. Though final value of f_D is selected based on visual analysis, f_D for any instance can be calculated as n/N where n is the number for data objects in proximity area and N is the total number of data objects in data set. For data set D45 and n = 1, f_D would be 0.022 and corresponding knn = 1, $r_{Outlier_Spatium} = 4.2168$, and #outliers = 13. Outliers will be those data objects which have less than knn data objects in their proximity area. This scenario is depicted in Fig. 1a in which outliers are the data objects which have less than knn i.e. 1 data object in their proximity area. Similarly, Fig. 1b illustrates the scenario for n = 2 and corresponding f_D = 0.044, knn = 2, and $r_{Outlier_Spatium} = 5.96734$. It can be observed that d_1 and $d\text{-}_2$ are considered outliers because they have less than 2 data objects in their proximity area. With same logic, d_3 is not an outlier any more. For Fig. 1f, f_D is 0.444 so if 44.4% of data objects are not in proximity area of a data object, then such data object will be considered outlier. For higher values of f_D, most of the data objects are gradually considered to be outlier because a higher percentage of data is expected to be in proximity area. By visually analyzing Fig. 1 and Table 1, it is observed that one can take any value of f_D between 0.0222 and 0.400 as number of outliers for the before mentioned range of f_D remain consistent i.e. number of outliers varies just between 10 to 14.

3.3 Steps of Outlier Identification

Step 1. Based on visual analysis, assume a fraction of data(f_D) and compute knn.
Step 2. Using Computed knn and size of data, $r_{Outlier_Spatium}$ is calculated as per Ester (1996).
Step 3. Using $r_{Outlier_Spatium}$ and Euclidean distance, proximity area is computed in 2D(circle) or 3D(sphere) space.
Step 4. Considering each data object (d_j) as center of proximity area, count number of data objects (C_j) in its respective proximity area.
Step 5. Data object d_j is an outlier if C_j is less than knn (computed in Step 1).

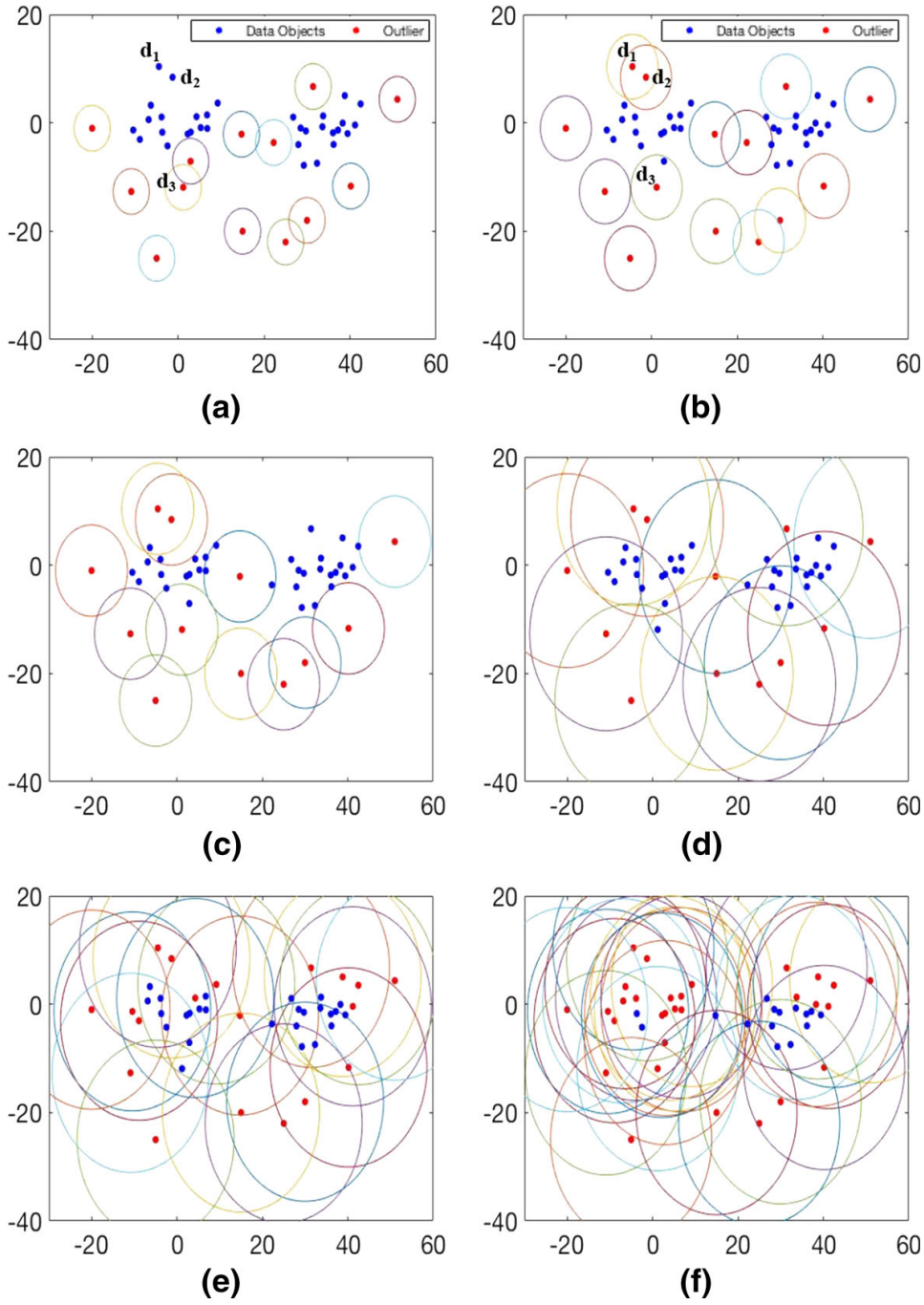


Fig. 1 Illustration of working logic of outlier identification by taking various values of f_D for D45 Data Set

DB-Outlier algorithm
Input: set of data objects(D), fraction of data(f_D)
Output: Outlier Matrix (OM), $Count_j$: count of data objects in otspatium of d_j
Outlier Identification:
Step 1. Set OM=ϕ; $Count_j$ =1 for ∀j ϵ 1,2,..,size(D).
Step 2. Calculate knn, using equation (11).
Step 3. Calculate $r_{otspatium}$ ← otspatium as per Ester et al. [21].
Step 4. For ∀j ϵ 1,2,..,size(D)
Begin
For ∀i ϵ 1,2,..,size(D) and j≠i
Begin
$dist_{ij} \leftarrow dist(d_i, d_j)$
If($dist_{ij} \leq r_{otspatium}$)
Begin
$Count_j$++
EndIf
End
End
Step 5. For ∀j ϵ 1,2,..,size(D)
Begin
If($Count_j$<knn)
Begin
OM←OM U $\{d_j\}$
End
End
Step 6. Return Count and OM

Time Complexity Analysis:

$$\Rightarrow T\,(C1\Theta\,(1) + C2\Theta\,(1) + C3\Theta\,(1))$$
$$+ T\,(C4\Theta\,(n+1)\,(C5\Theta\,(n)\,(C6\Theta\,(1) + C7\Theta\,(1) + C8\Theta\,(1))))$$
$$+ T\,(C9\Theta\,(n)\,(C10\Theta\,(1) + C11\Theta\,(1)))$$

On simplification:

$$\Rightarrow T(C123\Theta(1)) + T(C4\Theta(n+1)(C5\Theta(n)(C678\Theta(1)))) + T(C9\Theta(n)(C1011\Theta(1)))$$
$$\Rightarrow T(C_{123}\Theta(1)) + T(C4\Theta(n+1)(C5\Theta(n)(C678\Theta(1)))) + T(C9\Theta(n)(C1011\Theta(1)))$$
$$\Rightarrow T(C123\Theta(1) + C4\Theta(n+1)(C5\Theta(n)(C678\Theta(1))) + C9\Theta(n)(C1011\Theta(1)))$$
$$\Rightarrow T(C123\Theta(1) + C4\Theta(n+1)(C5678\Theta(1)) + C91011\Theta(n))$$
$$\Rightarrow T\left(C123\Theta(1) + C45678\Theta(n) + C45678\Theta(1) + C491011\Theta\left(n^2\right)\right)$$
$$\Rightarrow \Theta\left(n^2\right)$$

where C123 = C1 + C2 + C3, C678 = C6 + C7 + C8, C1011 = C10 + C11, C5678 = C5*C678, C91011 = C9*C1011, C45678 = C4*C5678, C491011 = C4*C91011.

Here, C1, C2,…, C11 are integer constants and n is size of data set, i.e. n = ||D||. Θ(1) is asymptotic notation representing a constant time taken for the respective operation. As Step 1, Step 2, and Step 3 takes constant time for execution, so their time complexity is Θ(1). Each step is performing different operation so different integer constants C1, C2, and C3 are taken as multiplication factor for time complexity analysis of each respective step. Thus, time complexity for first three steps is T(C1Θ(1) + C2Θ(1) + C3Θ(1)), which can be simplifies as C123Θ(1). Now, time complexity analysis for step 4 is $T\,(C4\Theta(n+1)(C5\Theta(n)(C6\Theta(1) + C7\Theta(1) + C8\Theta(1))))$. In step 4, outer loop body will run for n times whereas inner loop body will run only for n − 1 times. Additionally, there will be one extra iteration in any loop accounting for loop failure condition. Thus, time complexity

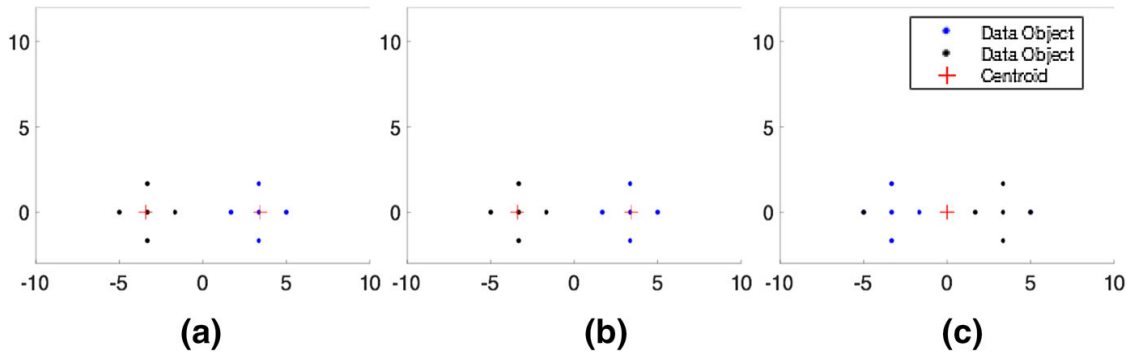


Fig. 2 Clustering Results of KIFCM algorithm using various kernel on D10 Data-set **a** Gaussian Kernal KIFCM, **b** Radial Basis Kernel KIFCM, **c** Hyper Tangent Kernal KIFCM

for step 4 is $T(C4\Theta(n+1)(C5\Theta(n)(C6\Theta(1)+C7\Theta(1)+C8\Theta(1))))$. Similarly, time complexity for step 5 is $T(C9\Theta(n)(C10\Theta(1)+C11\Theta(1)))$. On simplification, Time Complexity of DB-Outlier algorithm is $\theta\left(||D||^2\right)$.

3.4 Kernel Process

KIFCM-Gaussian is an extension of IFCM by exploiting Gaussian kernel function [1] for distance computation between any data object and cluster centroids. Kernel function maps input space to high dimensional space called as feature space [1]. Various kernel functions have been proposed in literature, most often used kernel function are [32]:

1. Polynomial Kernel [25]: $K(x_i, x_j) = \left(x_i.x_j + c\right)^d$, where c >=0 and d ϵ N.
2. Gaussian Kernel [25]: $K(x_i, x_j) = \exp\left(\frac{\left|x_{ii}^a - x_{jj}^a\right|^b}{\sigma^2}\right)$, where $\sigma, a, b > 0$
3. Hyper Tangent Kernel [15]: $K(x_i, x_j) = 1 - \tanh\left(-\left(\left\|x_i - x_j\right\|^2 \Big/ \sigma^2\right)\right), \quad \sigma > 0$
4. Radial Basis Kernel [15]: $K(x_i, x_j) = \exp\left(\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$, where $\sigma > 0$

The proposed work increases the accuracy of the IFCM by modifying the distance calculation between data objects and cluster centers through exploitation of Gaussian kernel function. Using many data sets, it was found heuristically that Gaussian kernelized IFCM gives best results. Figure 2 shows the results of these kernelized IFCM algorithm on D10 Data set.

3.5 Clustering Process

We are constructing the Distance Based Approach of Gaussian kernelized intuitionistic Fuzzy C-Means. Objective function of proposed algorithm is:

$$J_{DBKIFCM} = \sum_{i=1}^{nc+1}\sum_{j=1}^{n} md_{ji}^{*m}\left\|\Phi\left(d_j\right) - \Phi(C_i)\right\|^2 + \sum_{i=1}^{nc+1}\pi_i^* e^{1-\pi_i^*} \tag{15}$$

where d_j is j^{th} data object of given Data set $D = \{d_1, d_2, d_3, \ldots, d_n\}$; C_i is i^{th} centroid; md_{ji} is fuzzy membership degree and $md_{ji}^* = md_{ji} + \pi_{ji}$ is IFCM membership degree of d_j in

C_i; $dist_{ji}$ is Euclidean distance between d_j and C_i; nc is count of clusters; n is count of data objects; Π_{ji} is hesitation degree [1], defined as follows:

$$\pi_{ji} = 1 - md_{ji} - \left(1 - md_{ji}^{\alpha}\right)^{\frac{1}{\alpha}} \quad \text{with } \alpha > 0 \tag{16}$$

$$\pi_i^* = \frac{1}{n}\sum_{j=1}^{n} \pi_{ji} \; j \in [1, n] \tag{17}$$

and $\Phi(d_j) - \Phi(C_i)$ denotes feature space distance between C_i and d_j. The distance calculation in feature space via using a kernel in the input space is as follows [23]:

$$\begin{aligned} \Phi_{dist_{ji}^2} &= \left\|\Phi\left(d_j\right) - \Phi(C_i)\right\|^2 \\ &= \left(\Phi\left(d_j\right) - \Phi(C_i)\right) \cdot \left(\Phi\left(d_j\right) - \Phi(C_i)\right) \\ &= \Phi\left(d_j\right) \cdot \Phi\left(d_j\right) - 2\Phi\left(d_j\right) \cdot \Phi(C_i) + \Phi(C_i) \cdot \Phi(C_i) \\ &= K\left(d_j, d_j\right) - 2K\left(d_j, C_i\right) + K(C_i, C_i) \end{aligned} \tag{18}$$

As in the proposed technique, we adopted Gaussian Kernel [18]:

$$K(x, y) = \exp\left(\frac{|x^a - y^a|^b}{h^2}\right) \quad \text{with } a, b > 0 \tag{19}$$

where a and b are data specific and h is a positive number and defines kernel width, then K(x,x) = 1. Hence, Eq. (18) can be rewritten as:

$$\left\|\Phi\left(d_j\right) - \Phi(C_i)\right\|^2 = 2\left(1 - K\left(d_j, C_i\right)\right) \tag{20}$$

And Eq. (15) can be reformulated as:

$$J_{DBKIFCM} = 2\sum_{i=1}^{nc+1}\sum_{j=1}^{n} md_{ji}^{*m}\left(1 - K\left(d_j, C_i\right)\right) + \sum_{i=1}^{nc+1} \pi_i^* e^{1-\pi_i^*} \tag{21}$$

$$J_{DBKIFCM} = 2\sum_{i=1}^{nc+1}\sum_{j=1}^{n} md_{ji}^{*m}\left(1 - e^{-\frac{\left(d_j^a - C_i^a\right)^b}{h^2}}\right) + \sum_{i=1}^{nc+1} \pi_i^* e^{1-\pi_i^*} \tag{22}$$

For a given set of data objects D, we minimize $J_{DBKIFCM}$ in order to determine md_{ji}^* and C_i, subject to the following condition:

$$\sum_{i=1}^{nc} md_{ji}^* = 1 \quad 0 \le md_{ji}^* \le 1; \quad for\ all \quad j = 1, 2, 3, \ldots, n \tag{23}$$

and $md_{j(nc+1)}^* = 1$ when j denoting an outlier and $\sum_{i=1}^{nc} md_{ji}^* = 1$ for j denoting any data object.

Theorem *To find optimal clusters, it is a necessary condition to minimize $J_{DBKIFCM}$ subject to the condition specified in Eq. (23).*

On minimizing (15) under condition mentioned in (23), we get:

$$md_{ji}^* = \frac{1}{\sum_{j=1}^{nc+1}\left(\frac{\Phi_{d_{ic}^2}}{\Phi_{d_{ij}^2}}\right)^{\frac{1}{m-1}}} \tag{24}$$

and centroids are calculated as:

$$C_i = \frac{\sum_{j=1}^{n} md_{ji}^{*m}\left(d_j\right)}{\sum_{j=1}^{n} md_{ji}^{*m}} \tag{25}$$

Using DB-Outlier algorithm, we identify outliers and assign zero to md_{ji}^{*} *to data objects identified as outliers. Thus, modified membership equation as per DBKIFCM is as*:

$$md_{ji}^{*} = \begin{cases} \dfrac{1}{\sum_{j=1}^{nc+1}\left(\frac{\Phi_{d_{ic}^2}}{\Phi_{d_{ij}^2}}\right)^{\frac{1}{m-1}}} & \text{if } \left(\left\|d_j \middle| dist(d_i,d_j) \le r_{Outlier_Spatium}\right\|\right) \ge \text{knn} \quad \text{for} \quad \{\forall j \in [1..N],\ j \ne i\} \\ 0 \quad \text{if } \left(\left\|d_j \middle| dist(d_i,d_j) \le r_{Outlier_Spatium}\right\|\right) < \text{knn} \quad \text{for} \quad \{\forall j \in [1..N],\ j \ne i\} \end{cases} \tag{26}$$

Proof Differentiate $J_{DBKIFCM}$ wrt md_{ji}^{*} and C_i and the derivatives are set to zero. Thus, we get Eqs. (40) and (47). Step by step proof is given in Appendix A.

DBKIFCM Algorithm

Input: set of data objects (D), Cluster count (nc+1), Maximum Iteration Count (max_iter), Minimum improvement (SC), Fuzziness index (m), Fraction of Data(f_D)
Output: Centroid matrix (C), Outlier matrix (OM)
Clustering process:

Step 1. [OM, Count] ← DB_Outlier(D, f_D)

Step 2. Set hyperparameters like m, max_iter, SC

Step 3. Set initial centroids C_1, C_2, …., C_{nc} to d_l, d_m, …., d_z such that d_l, d_m, …., d_z has maximum value for Count, such that $Count_l \le Count_m \le Count_n \le \ldots \le Count_z$

Step 4. Set md_{ij}^{*}, ∀j ϵ [1,2,..,n], subject to constraints in equation (23) using (42)

Step 5. For ∀ d_k ϵ OM, set $md_{ki} = 0$ ∀i

```
For k=1, 2, .. , max_iter; do:
Begin
    Using (49), update C
    Using (22), compute Obj_fun_k
    Using (42), update md_ji^* ∀j
    If (|Obj_fun_k - Obj_fun_k-1|)< SC
        Begin
            Break
        EndIf
End
End
```

Step 6. Return C and md

Time Complexity Analysis:

$$\Rightarrow T\left(C1\Theta\left(n^2\right) + C2\Theta(1) + C3\Theta(n\log n) + C4\Theta(n)\right)$$
$$+ T(C5\Theta(1) + C6\Theta(\max_iter).(C7\Theta(n) + C8\Theta(nc.n) + C9\Theta(nc) + C10\Theta(1)))$$
$$+ T(C11\Theta(1))$$

$$\Rightarrow T\left(C1\Theta\left(n^2\right) + C2\Theta(1) + C3\Theta(n\log n) + C4\Theta(n)\right)$$
$$+ T(C5\Theta(1) + C67\Theta(\max_iter.n) + C68\Theta(\max_iter.nc.n) + C69\Theta(\max_iter.nc) + C610\Theta(\max_iter))$$
$$+ T(C11\Theta(1))$$

$$\Rightarrow T\left(\begin{array}{r} C1\Theta\left(n^2\right) + C2\Theta(1) + C3\Theta(n\log n) + C4\Theta(n)+ \\ (C5\Theta(1) + C67\Theta(\max_iter.n) + C68\Theta(\max_iter.nc.n) + C69\Theta(\max_iter.nc) + C610\Theta(\max_iter))+ \\ (C11\Theta(1))\end{array}\right)$$

$$\Rightarrow T\left(C1\Theta\left(n^2\right) + C68\Theta(\max_iter.nc.n)\right) = \Theta\left(n^2 + \max_iter.n\right)$$

$Hence$, $\Theta\left(n^2 + \max_iter.n\right)$ is time complexity of DBKIFCM.

where C67 = C6*C7, C68 = C6*C8, C69 = C6*C9, C5610 = C6*C10

Similar to time complexity analysis of DB_Outlier Algorithm, C1, C2,…, C11 are integer constants and n is size of data set, i.e. n = ||D||. $\Theta(1)$ is asymptotic notation representing a constant time taken for the respective operation. As per time complexity calculation mentioned in Sect. 3.2, time complexity DB_Outlier is $\Theta\left(n^2\right)$. Step 2 and Step 4 take constant time for execution, so their time complexity is $\Theta(1)$. Step 3 does sorting on Count_i using merge sort, so its time complexity is $\Theta(\log n)$. Thus, $T\left(C1\Theta\left(n^2\right) + C2\Theta(1) + C3\Theta(n\log n) + C4\Theta(n)\right)$ represents total time complexity from Step 1 to Step 4. For Step 5, time complexity is T $(C5\Theta(1) + C6\Theta(\max_iter).(C7\Theta(n) + C9\Theta(nc) + C10\Theta(1)))$ as outer loop runs for iterations equal to number of outliers, which will be constant for a given data set. Inner loop will be executed max_$iter$ times, resulting in total time complexity of inner loop as $C6\Theta(\max_iter).(C7\Theta(n) + C8\Theta(nc.n) + C9\Theta(nc) + C10\Theta(1))$. For simplification, we consider nc to be constant for a given data set and resultant time Complexity of DBKIFCM will be $\Theta\left(\|D\|^2 + \max_iter.\|D\|\right)$.

4 Simulations and Results

For experimentation, simulation is done on a system with MATLAB 2017a, intel i5 processor of 2.5 GHz and 8 GB RAM. Experimental results are projected using tables and figures to provide a comparative view of existing algorithms with the proposed algorithm DBKIFCM. For experimental simulation, we have considered following parameters as: fuzziness index(m) equal to 2, maximum iterations (SC) to be 100 and minimum relative improvement in consecutive iterations is 1-e5. We used Gaussian kernel in DBKIFCM. Parameters a, b and sigma (kernel width) of Gaussian kernel have a significant effect on the results of algorithm. But choice of appropriate value for them is an "open problem". Thus, we adopted "hit and trial" method to reach the appropriate values for a, b, and sigma. Experiments are performed on synthetic 2D data sets and real-world data set. Notations used to represent clustering results are: Green/Blue/Black dot for different clusters, '+' for centroids, and '○' for outlier.

4.1 Synthetic Data Sets

To validate and conclude our findings and salient features of our proposal, we experimented on a number of popular data sets with different data characteristics such as uniformly and non-uniformly spread data, linearly separable and non-linearly separable data, data with

Table 2 Brief description of data sets

Data Set	||D||	No. of clusters	Number of outliers	Remarks
D12	12	2	1	Uniformly Distributed Data Objects with an outlier and a noise point
D15	15	2	4	Uniformly Distributed Data Objects with equal size clusters and few outliers
DUNN	138	2	10–15	Density variant clusters

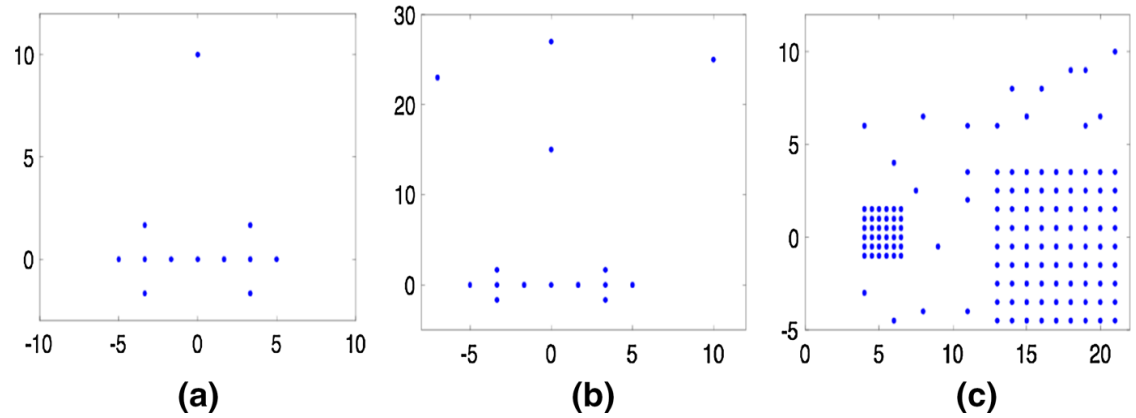


Fig. 3 **a** D12 data set, **b** D15 Data Set, **c** Dunn Data Set

Table 3 Brief description of real data sets

Name of Data Set	||D||	No. of clusters	Dimensionality of data set	Remarks
Fisher-iris	150	3	4	3 species of 50 flowers each and one species can be separated from other two based on linear discriminant model, but latter two cannot be.
Breast Cancer Wisconsin	569	2	32	32 features calculated from each digital image of abreast mass.
Wine	178	3	13	Wine samples from 3 cultivators from a region in Italy.

highly unlike cluster size, density variant data set, and size variant cluster data sets. Worked data sets are listed in Table 2 and plot is shown in Fig. 3.

4.2 Real Data Sets

Proposed approach is also tested on real world high dimensional data sets. Table 3 lists some few popular real data sets along with their salient features that we considered for experimental analysis.

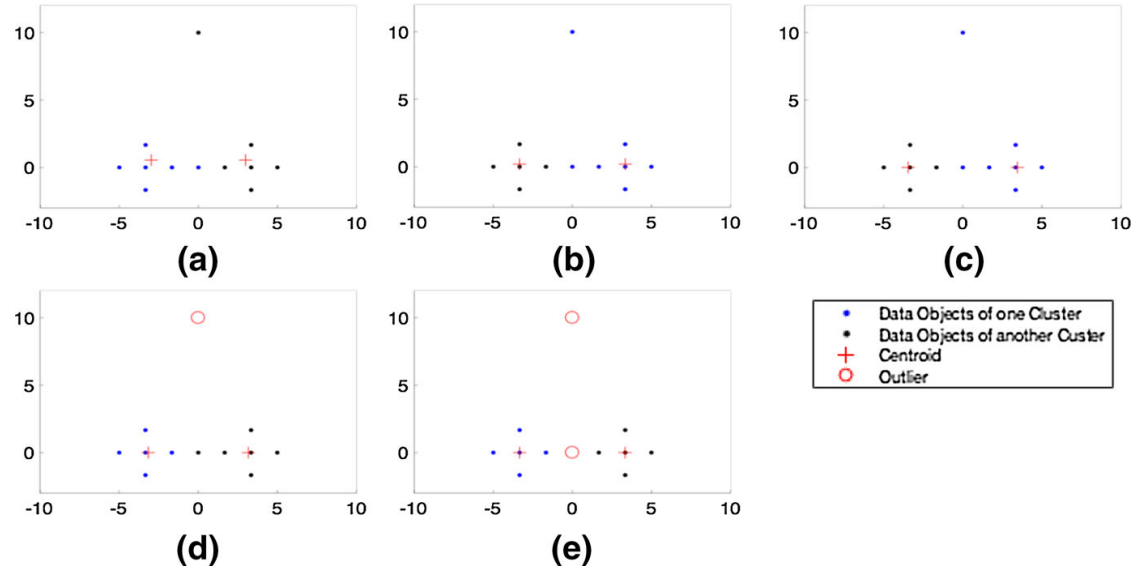


Fig. 4 D12 Data Set: Clustering results of **a** FCM, **b** IFCM, **c** KIFM(RBF), **d** DOFCM, **e** DBKIFM (Gaussian)

Table 4 Computed Centroids and average error on D12 by FCM, IFCM, KIFCM(RBF), DOFCM and DBKIFCM

	Cluster 1		Cluster 2		Average error
	cx*	cy*	cx*	cy*	
FCM	2.985473	0.54363	−2.98527	0.543426	0.421186
IFCM	3.34448	0.190049	−3.34447	0.190096	0.036148
KIFCM(RBF)	3.455145	0.00106	−3.45421	0.001056	0.013153
DOFCM	3.167435	3.68E−08	−3.16724	−3.56E−08	0.029812
DBKIFCM(Gaussian)	3.330091	6.67E−18	−3.33393	−4.8E−20	6.75E−05

4.3 Comparing Results Using Synthetic Data

Example 1 Comparison on D12 and D15 data sets: D12 data set consists of 2 clusters with 5 data objects in each cluster (d_1 to d_{10}), a noise point: d_{11} and an outlier: d_{12}. Similarly, D15 consists of 2 clusters with 5 data objects in each cluster (d_1 to d_{10}), a noise point: d_{11} and 4 outliers (d_{12} to d_{15}). The ideal centroids for D12 and D15 data sets are:

$$C_{ideal} = \begin{bmatrix} 3.34 & 0 \\ -3.34 & 0 \end{bmatrix}$$

Figure 4 and Table 4 draw a comparison of clustering results of FCM, IFCM, KIFCM(RBF), DOFCM and DBKIFCM on D12 data set. Table 4 lists centroid coordinates and error in computing ideal centroids. The error is computed as:

$$E^{\#} = \|C_{ideal} - C_{\#}\|^2 \tag{27}$$

where # is FCM/IFCM/KIFCM/DOFCM/DBKIFCM.

It can be observed from Fig. 4 that FCM, IFCM, and KIFCM fail to identify the presence of outlier and noise, and treats them like a normal data object. Due to which, error in centroid computation is high and Table 4 validates it by showing approx. 42.1% error in centroid computation by FCM.

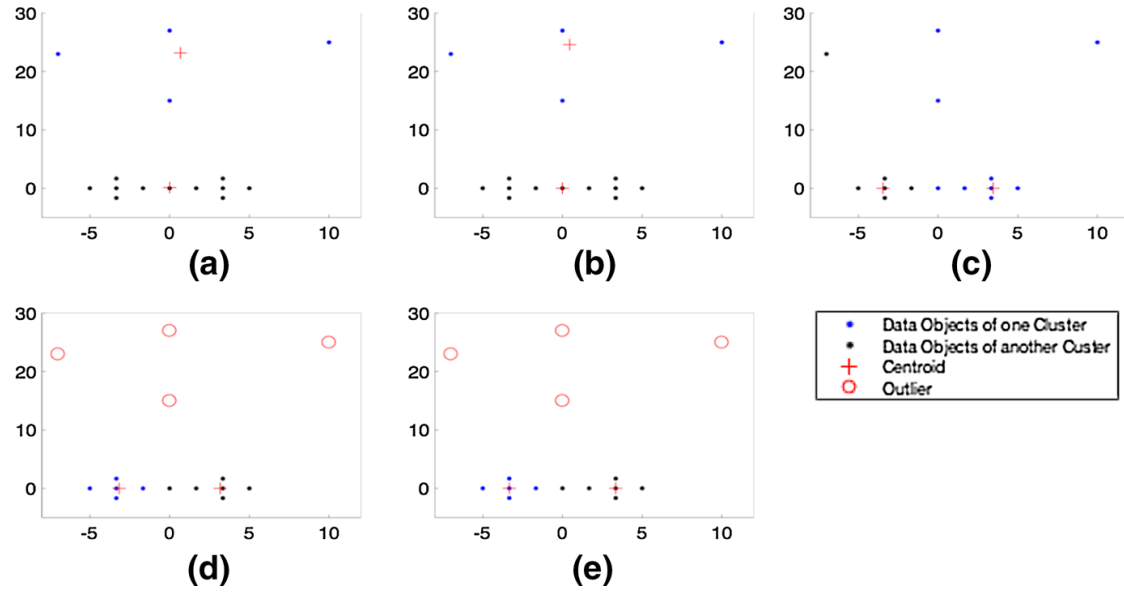


Fig. 5 D15 Data Set: Clustering results of **a** FCM, **b** IFCM, **c** KIFM(RBF), **d** DOFCM, **e** DBKIFM (Gaussian)

Table 5 Computed centroids and average error on D15 by FCM, IFCM, KIFCM, DOFCM, and DBKIFCM

	Cluster 1		Cluster 2		Average error
	cx*	cy*	cx*	cy*	
FCM	0.004743942	0.122739029	0.675725149	23.17390526	282.1474656
IFCM	−0.004502973	0.000753444	0.455680647	24.60629122	315.5312298
KIFCM (RBF)	3.45434884	4.87727E−05	−3.455664299	5.04816E−05	0.013226946
DOFCM	3.167428882	−1.82028E−08	−3.167250349	1.73372E−08	0.029811616
DBKIFCM (Gaussian)	3.340216448	1.62905E−18	−3.340216448	1.31837E−20	4.68496E−08

IFCM performs much better than FCM in computing centroid but still fails to identify outliers and noise. KIFCM(RBF) reduces the effect of outlier and noise, improving error in centroid computation. DOFCM identifies the outliers and improves accuracy of centroid computation over FCM and IFCM. DBKIFCM outperforms all these algorithms in identifying outlier, noise and improving centroid accuracy.

Figure 5 and Table 5 draw a comparison of clustering results of FCM, IFCM, KIFCM(RBF), DOFCM and DBKIFCM on D15 data set. Table 5 lists the centroid coordinates and error in computing ideal centroid on D15 data set. Error is computed as per Eq. (24). In D15 data set, 26.67% of data objects are outliers. It can be observed from Fig. 5 and Table 5 that FCM and IFCM gets totally confused between actual clusters and outliers, resulting one cluster of all outliers and other of all data objects and noise. Hence, cluster accuracy is very poor. KIFCM(RBF) performance much better than FCM and IFCM in terms of resultant clusters and centroids accuracy. But, it fails to identify outliers which are identified by DOFCM and DBKIFCM. DOFCM succeeds in identifying outliers but its centroid accuracy is lower than KIFCM. DBKIFCM outperforms all these algorithms in outlier identification, noise reduction and centroid computation.

Example 2 Comparison on Dunn data set: Dunn data is also known as 2D Square data set with 2 square shaped clusters of varying density. Figure 3c shows Dunn data set with noise, and

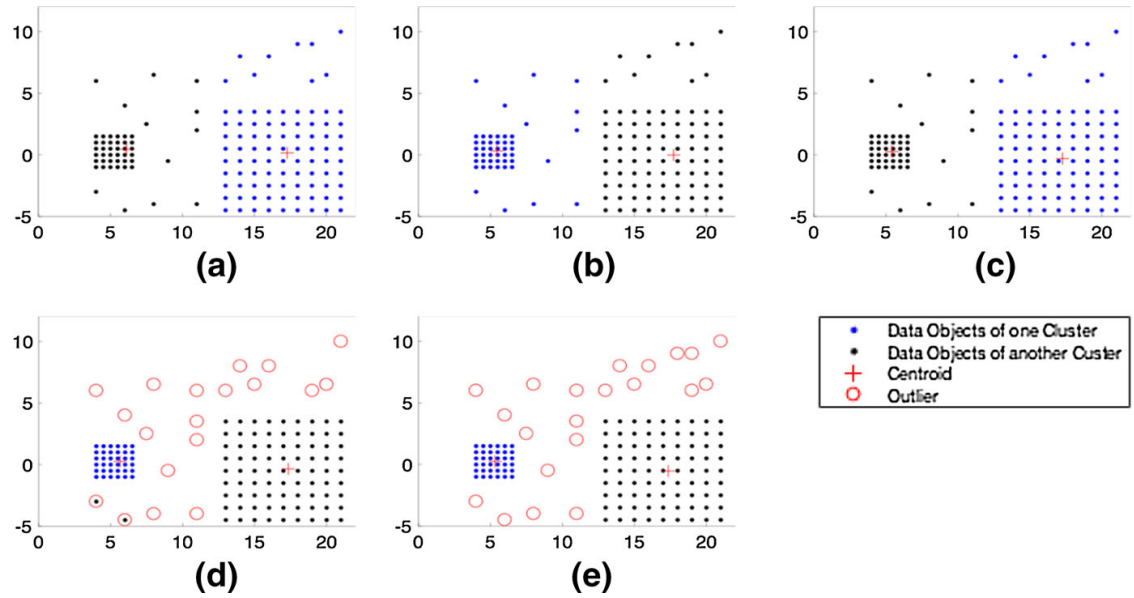


Fig. 6 **a** Noisy DUNN data set, **b** FCM clustering results, **c** KIFCM clustering results, **d** IFCM clustering results, **e** DBKIFCM clustering results, **f** DOFCM clustering results

Table 6 Computed centroids and average error on DUNN by FCM, IFCM, KIFCM, DOFCM, and DBKIFCM

	Cluster 1		Cluster 2		Average error
	cx*	cy*	cx*	cy*	
FCM	17.27373276	0.159051402252611	6.118821026	0.455801258635604	0.653241253
IFCM	17.72622645	−0.00960325619946819	5.453973291	0.324924797430614	0.407556323
KIFCM (RBF)	17.26267026	−0.298513851449195	5.410559755	0.276345555277432	0.068032928
DOFCM	17.33611773	−0.342985956159516	5.656627539	0.257531267551569	0.151515608
DBKIFCM (Gaussian)	17.34706059	−0.524737063614454	5.316894807	0.246825892504632	0.062773982

Fig. 6 and Table 6 draws comparison of clustering results of FCM, IFCM, KIFCM(RBF), DOFCM and DBKIFCM. Table 6 lists centroid coordinates and error in computing ideal centroids. Error is computed as per Eq. (27). The ideal centroids for Dunn data set are:

$$C_{ideal} = \begin{bmatrix} 17 & -0.5 \\ 5.25 & 0.25 \end{bmatrix}$$

It can be observed from Fig. 6 and Table 6 that FCM and IFCM treats noise and outliers as normal data objects, resulting in decent clusters. KIFCM(RBF) results in more accurate centroids. DOFCM identifies all outliers and results in more accurate centroids than FCM and IFCM. However, it assigns 2 outliers to a cluster as well. DBKIFCM outperforms FCM, IFCM, KIFCM(RBF) and DOFCM.

Table 7 Misclassification count and accuracy for FCM, IFCM, KIFCM(RBF), DOFCM and DBKIFCM

Algorithm	Fisheriris		Wisconsin Breast Cancer		Wine	
	Misclassification	Accuracy	Misclassification	Accuracy	Misclassification	Accuracy
FCM	16	0.89333	83	0.8541	56	0.6854
IFCM	14	0.90000	75	0.6467	54	0.6966
KIFCM	11	0.71333	53	0.9080	47	0.7360
DOFCM	16	0.89333	51	0.9000	49	0.7370
DBKIFCM	7	0.92000	52	0.9086	46	0.7350

4.4 Comparison of DBKIFCM with Existing Algorithms Using Real Data Sets

Example 3 High Dimensional Real World Data Set: Fisheriris, Wisconsin Breast Cancer and Wine are briefly described in Table 3. For high dimensional data sets, results are evaluated using Huang's accuracy measure [1] as follows:

$$r = \frac{\sum_{i=1}^{nc} n_i}{n} \tag{28}$$

where n = ||D|| and n_i is count of correctly clustered data objects of ith cluster. Value of r ranges in [0,1], and for perfect clustering, value of r will be equal to 1. Table 7, lists misclassification and the accuracy of FCM, IFCM, KIFCM(RBF), DOFCM and DBKIFCM clustering algorithms for High Dimensional Real-World data set. It can be observed that DBKIFCM outperforms all other.

4.5 Sensitivity to Initialization (STI)

STI is one of major problems of core fuzzy clustering algorithm (FCM). As shown in Fig. 7, it is observable that final clustering results are greatly affected by the choice of initial centroids on data set D1 [7]. If initial clusters are chosen to be $\begin{bmatrix} -0.5 & 1.7 \\ 0.4 & 2.0 \\ 0.0 & 2.7 \end{bmatrix}$ as shown in Fig. 7a, then resultant clusters after 12 iteration are as shown in Fig. 7b. But if initial centroids are chosen to be $\begin{bmatrix} 0.2 & -0.7 \\ -0.1 & 3.0 \\ 0.6 & 1.9 \end{bmatrix}$ as shown in Fig. 7c, then resultant clusters after 12 iteration are as shown in Fig. 7d and these resultant clusters are inaccurate. This problem can be solved by using DBKIFM to determine initial centroids instead of choosing random centroid. However, DBKIFM performs well for data sets that are either nearly uniform distributed or have equal cluster density. DBKIFM identify outliers first and then initializes centroids as those data objects which have highest count of data objects in their respective proximity area. Using DBKIFCM, problem of STI can be solved and iteration count to produce resulting clusters is effectively improved. This is illustrated using Table 8.

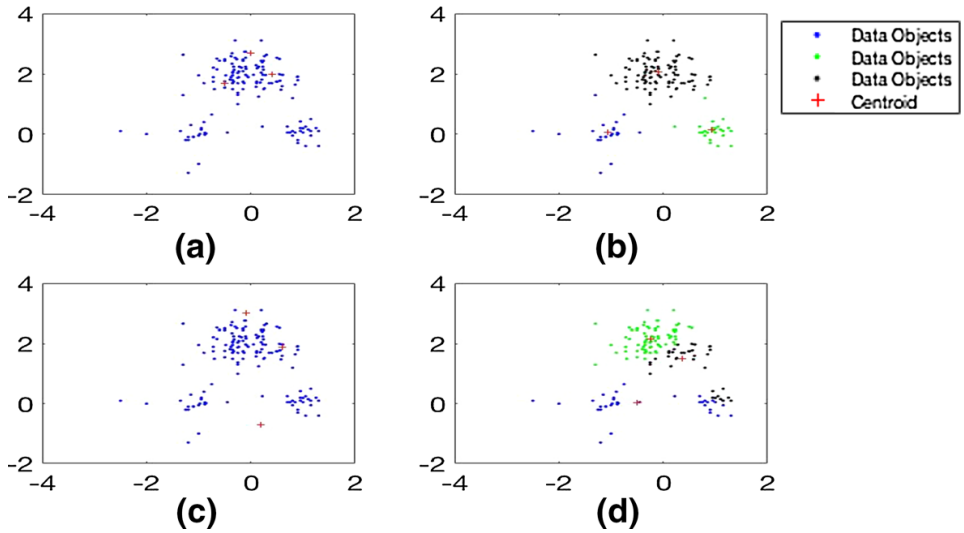


Fig. 7 Sensitivity to Initialization on D1

Table 8 Iteration Count for KIFCM (RBF) and DBKIFCM on different data sets

Data Set	No. of Iterations KIFCM(RBF)	No. of Iterations DBKIFCM
D1	32–63	28
D12	8–73	12
D15	12–100	49
Dunn	19–100	25

5 Conclusion

Clustering is widely applied across many real life problems such as image classification, object recognition etc. Various clustering algorithms have been proposed in literature with varied success. To overcome the problem of STI, segment non-linear data effectively and make clustering process more robust to the presence of noise and outliers, we have proposed a new clustering algorithm named as DBKIFCM. It is used to identify outliers, initializes centroids and then performs clustering. Distance Based approach is used to identify outliers and Gaussian Kernel function is used to map data objects to feature space in order to improve 'resulting clusters accuracy'. IFCM is used to perform clustering in feature space. This proposed algorithm has significantly outperformed earlier proposed algorithms in computing final clusters and centroids position. Performance of DBKIFCM is tested on synthetic data sets in literature such as Diamond (D12 and D15), and Dunn-Square and real data set such as Fisheriris, Wisconsin Breast Cancer, and Wine. A comparison of proposed approach is drawn with FCM, IFCM, KIFCM(RBF) and DOFCM. It is also found that DBKIFCM has overcome the problem of STI and is significantly effective in the presence of outliers and noise as well.

Appendix A: Distance Based Approach of Gaussian Kernelized Intuitionistic Fuzzy C Means

Proof of Distance-Based Approach of Gaussian Kernelized Intuitionistic Fuzzy C Means

Proof of DBKIFCM which is distance-based approach of KIFCM is given in this appendix. Minimization of Eq. (15), subjected to constraint in Eq. (23) is done by minimizing the following function:

$$J_{DBKIFCM} = \sum_{i=1}^{nc+1}\sum_{j=1}^{n} md_{ji}^{*m}\left\|\Phi\left(d_j\right) - \Phi(C_i)\right\|^2 + \sum_{i=1}^{nc+1}\pi_i^* e^{1-\pi_i^*} + \sum_{i=1}^{nc+1}\lambda_i\left(\sum_{j=1}^{n} md_{ji}^* - 1\right) \tag{29}$$

where λ_i (i = 1, 2, 3,…, nc) are Langrangian multipliers.

So, to solve this problem, we need to specify $\left\|\Phi\left(d_j\right) - \Phi(C_i)\right\|^2$ as follows:

$$\begin{aligned}\left\|\Phi\left(d_j\right) - \Phi(C_i)\right\|^2 &= \left(\Phi\left(d_j\right) - \Phi(C_i)\right).\left(\Phi\left(d_j\right) - \Phi(C_i)\right)\\ &= \Phi\left(d_j\right).\Phi\left(d_j\right) - 2\Phi\left(d_j\right).\Phi(C_i) + \Phi(C_i).\Phi(C_i)\\ &= K\left(d_j, d_j\right) - 2K\left(d_j, C_i\right) + K(C_i, C_i)\end{aligned}$$

As in the proposed technique, we adopted Gaussian Kernel, that is:

$$K(x, y) = \exp\left(-\frac{\sum|x_i^a - y_i^a|^b}{h^2}\right) \tag{30}$$

where h denotes kernel width, and a, b are data specific positive number, then K(x,x) = 1. Hence,

$$\left\|\Phi\left(d_j\right) - \Phi(C_i)\right\|^2 = 2\left(1 - K\left(d_j, C_i\right)\right) \tag{31}$$

So, Eq. (26) modifies to:

$$J_{DBKIFCM} = 2\sum_{i=1}^{nc+1}\sum_{j=1}^{n} md_{ji}^{*m}\left(1 - K\left(d_j, C_i\right)\right) + \sum_{i=1}^{nc+1}\pi_i^* e^{1-\pi_i^*} + \sum_{i=1}^{nc+1}\lambda_i\left(\sum_{j=1}^{n} md_{ji}^* - 1\right) \tag{32}$$

Or

$$J_{DBKIFCM} = 2\sum_{i=1}^{nc+1}\sum_{j=1}^{n} md_{ji}^{*m}\left(1 - \exp\left(-\frac{\sum\left|d_j^a - C_i^a\right|^b}{h^2}\right)\right) + \sum_{i=1}^{nc+1}\pi_i^* e^{1-\pi_i^*} + \sum_{i=1}^{nc+1}\lambda_i\left(\sum_{j=1}^{n} md_{ji}^* - 1\right) \tag{33}$$

Partial Derivative of $J_{DBKIFCM}$ (Given in Eq. (30)) with Respect to md_{ji}^*

The partial derivative of $J_{DBKIFCM}$ wrt md_{ji}^* is:

$$\frac{\partial J}{\partial md_{ji}^*} = 2m \cdot md_{ji}^{*(m-1)}\left(1 - K\left(d_j, C_i\right)\right) + \lambda_i \tag{34}$$

Equate (31) to zero to solve for md^*_{ji}:

$$\frac{\partial J}{\partial md^*_{ji}} = 0 \tag{35}$$

$$\Rightarrow 2m \cdot md^{*(m-1)}_{ji}\left(1 - K\left(d_j, C_i\right)\right) + \lambda_i = 0 \tag{36}$$

$$\Rightarrow md^*_{ji} = \left(-\frac{\lambda_i}{2m.\left(1-K\left(d_j,C_i\right)\right)}\right)^{\frac{1}{(m-1)}} \tag{37}$$

To fulfil the following constraint:

$$0 \leq md^*_{ji} \leq 1$$

and $\sum_{i=1}^{nc+1} md^*_{ji} = 1$ for $\forall j \in [1, n]$.

$md^*_{(nc+1)j} = 1$ when j denoting an outlier and $\sum_{i=1}^{nc} md^*_{ji} = 1$ for j denoting any data point.

$$md^*_{ji} = \frac{md^*_{ji}}{\sum_{i=1}^{nc+1} md^*_{ji}} \tag{38}$$

using Eq. (34), Eq. (35) can be written as:

$$md^*_{ji} = \frac{\left(-\frac{\lambda_i}{2m.\left(1-K\left(d_j,C_i\right)\right)}\right)^{\frac{1}{(m-1)}}}{\sum_{i=1}^{nc+1}\left(\left(-\frac{\lambda_i}{2m.\left(1-K\left(d_j,C_i\right)\right)}\right)^{\frac{1}{(m-1)}}\right)} \tag{39}$$

$$\Rightarrow md^*_{ji} = \frac{\left(-\frac{1}{\left(1-K\left(d_j,C_i\right)\right)}\right)^{\frac{1}{(m-1)}}}{\sum_{i=1}^{nc+1}\left(\left(-\frac{1}{\left(1-K\left(d_j,C_i\right)\right)}\right)^{\frac{1}{(m-1)}}\right)} \tag{40}$$

Using following equation:

$$\Phi_{d^2_{ji}} = \left\|\Phi\left(d_j\right) - \Phi(C_i)\right\|^2 \tag{41}$$

Equation (43) can be re written as:

$$md^*_{ji} = \frac{1}{\sum_{j=1}^{nc+1}\left(\frac{\Phi_{d^2_{ic}}}{\Phi_{d^2_{ij}}}\right)^{\frac{1}{m-1}}} \tag{42}$$

Partial Derivation of $J_{DBKIFCM}$ with Respect to C_i

The partial derivation of $J_{DBKIFCM}$ with respect to C_i is:

$$\frac{\partial J}{\partial C_i} = 2\sum_{j=1}^{n} md^{*m}_{ji}\left(-\exp\left(-\frac{\sum\left|d^a_j - C^a_i\right|^b}{h^2}\right)\right)\left(\frac{b(d_j - C_i)^{(b-1)}(-1)(aC^{(a-1)}_i)}{h^2}\right) \tag{43}$$

$$\Rightarrow \frac{\partial J}{\partial C_i} = 2ab\sum_{j=1}^{n} md^{*m}_{ji}\left(\exp\left(-\frac{\sum\left|d^a_j - C^a_i\right|^b}{h^2}\right)\right)\left(\frac{\left(d^a_j - C^a_i\right)^{(b-1)}(C^{(a-1)}_i)}{h^2}\right) \tag{44}$$

$$\Rightarrow \frac{\partial J}{\partial C_i} = 2ab\sum_{j=1}^{n} md^{*m}_{ji}\left(K\left(d_j, C_i\right)\right)\left(\frac{(d_j - C_i)^{(b-1)}(C^{(a-1)}_i)}{h^2}\right) \tag{45}$$

To solve for C_i Equate Eq. (44) to zero:

$$\frac{\partial J}{\partial C_i} = 0 \tag{46}$$

$$\Rightarrow 2ab \sum_{j=1}^{n} md_{ji}^{*m} \left(K\left(d_j, C_i\right)\right) \left(\frac{\left(d_j - C_i\right)^{(b-1)} \left(C_i^{(a-1)}\right)}{h^2} \right) = 0 \tag{47}$$

$$\Rightarrow \sum_{j=1}^{n} md_{ji}^{*m} \left(K\left(d_j, C_i\right)\right) \left(\frac{\left(d_j - C_i\right)^{(b-1)} \left(C_i^{(a-1)}\right)}{h^2} \right) = 0 \tag{48}$$

From above, either $C_i = 0$ or $C_i = d_j$ or $md_{ji}^{*} = 0$; md_{ji}^{*} is 0 when d_j is an outlier. $C_i = 0$ or $C_i = d_j$ is not giving updation of C_i. So, update C_i using following equation:

$$C_i = \frac{\sum_{j=1}^{n} md_{ji}^{*m} \left(d_j\right)}{\sum_{j=1}^{n} md_{ji}^{*m}} \tag{49}$$

References

1. Kaur P, Soni AK, Gosain A (2013) A robust kernelized intuitionistic fuzzy c-means clustering algorithm in segmentation of noisy medical images. Pattern Recognit Lett
2. Han J, Kamber M, Pei J (2006) Data mining, Southeast Asia edition: concepts and techniques. Morgan Kaufmann, San Francisco
3. Rodríguez Ramos A, Rivera-Torres PJ, da Silva Neto AJ, Llanes-Santiago O (2018) A hybrid algorithm for fault diagnosis using fuzzy clustering tools. Fuzzy Optim Model J 1(1):11–30
4. Yu X (2014) Blurred trace infrared image segmentation based on template approach and immune factor. Infrared Phys Technol 67:116–120
5. Yu X, Yuan X, Dong E, Ríha K (2016) Target extraction of banded blurred infrared images by immune dynamical algorithm with two-dimensional minimum distance immune field. Infrared Phys Technol 77:94–99
6. Meenakshi S, Suganthi M, Sureshkumar P (2019) Segmentation and boundary detection of fetal kidney images in second and third trimesters using kernel-based fuzzy clustering. J Med Syst 43(7):203
7. Zadeh LA (1965) Fuzzy sets. Inf Control 8(3):338–353
8. Dunn JC (1973) A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. J Cybern 3:32–57
9. Bezdek JC, Ehrlich R, Full W (1984) FCM: the fuzzy c-means clustering algorithm. Comput Geosci 10(2–3):191–203
10. Kaur P, Gupta P, Sharma P (2012) Review and comparison of kernel based fuzzy image segmentation techniques. Int J Intell Syst Appl 4(7):50
11. Jain AK, Murty NM, Flynn PJ (1999) Data clustering: a review. ACM Comput Surv (CSUR) 31(3):264–323
12. Gosain A, Dahiya S (2016) Performance analysis of various fuzzy clustering algorithms: a review. Procedia Comput Sci 79:100–111
13. Kim D-W, Lee KH, Lee D (2004) Fuzzy clustering of categorical data using fuzzy centroids. Pattern Recognit Lett 25(11):1263–1271
14. Pal NR, Pal K, Keller JM, Bezdek JC (2005) A possibilistic fuzzy c-means clustering algorithm. IEEE Trans Fuzzy Syst 13(4):517–530
15. Chintalapudi KK, Kam M (1998) The credibilistic fuzzy c means clustering algorithm. In: 1998 IEEE international conference on systems, man, and cybernetics, 1998, vol 2. IEEE, pp 2034–2039
16. Kaur P, Soni AK, Gosain A (2013) Robust kernelized approach to clustering by incorporating new distance measure. Eng Appl Artif Intell 26(2):833–847
17. Kaur P (2011) Robust intuitionistic fuzzy cmeans clustering for linearly and nonlinearly separable data. In: 2011 international conference on image information processing, 11/2011
18. Xu Z, Junjie W (2010) Intuitionistic fuzzy C-means clustering algorithms. J Syst Eng Electron 21(4):580–590
19. Guan S, Wang Y, Chen X (2018) An approach of single-crystal defect detection using x-ray orientation instrument. IEEE Trans Instrum Meas 68(10):4122–4134

20. Huang H, Meng F, Zhou S, Jiang F, Manogaran G (2019) Brain image segmentation based on FCM clustering algorithm and rough set. IEEE Access 7:12386–12396
21. Debakla M, Salem M, Djemal K, Benmeriem K (2019) Fuzzy farthest point first method for MRI brain image clustering. IET Image Process 13(13):2395–2400
22. Li L, Yang Z, Jiao L, Liu F, Liu X (2019) High-resolution SAR change detection based on ROI and SPP net. IEEE Access 7:177009–177022
23. Chaira T (2011) A novel intuitionistic fuzzy C means clustering algorithm and its application to medical images. Appl Soft Comput 11(2):1711–1717
24. Kaur P, Soni AK, Gosain A (2013) RETRACTED: a robust kernelized intuitionistic fuzzy c-means clustering algorithm in segmentation of noisy medical images. Pattern Recognit Lett 34:163–175
25. Tushir M, Srivastava S (2010) A new Kernelized hybrid c-mean clustering model with optimized parameters. Appl Soft Comput 10(2):381–389
26. Kaur P, Gosain A (2011) A density oriented fuzzy C-means clustering algorithm for recognising original cluster shapes from noisy data. Int J Innov Comput Appl 3(2):77–87
27. Kaur P, Gosain A (2010) Density-oriented approach to identify outliers and get noiseless clusters in fuzzy C-means. In: 2010 IEEE international conference on fuzzy systems (FUZZ). IEEE, pp 1–8
28. Ramos AR, Llanes-Santiago O, Bernal de Lázaro JM, Corona CC, Silva Neto AJ, Verdegay Galdeano JL (2017) A novel fault diagnosis scheme applying fuzzy clustering algorithms. Appl Soft Comput 58:605–619
29. Rodríguez-Ramos A, da Silva Neto AJ, Llanes-Santiago O (2019) A proposal of hybrid fuzzy clustering algorithm with application in condition monitoring of industrial processes. In: Uncertainty management with fuzzy and rough sets. Springer, Cham, pp 3–30
30. Rodríguez-Ramos A, da SilvaNeto AJ, Llanes-Santiago O (2018) An approach to fault diagnosis with online detection of novel faults using fuzzy clustering tools. Expert Syst Appl 113:200–212
31. Ester M, Kriegel H-P, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: Kdd, vol 96, no. 34, pp 226–231
32. Liu X, Zhu X, Li M, Wang L, Zhu E, Liu T, Kloft M, Shen D, Yin J, Gao W (2019) Multiple kernel k-means with incomplete kernels. IEEE Trans Pattern Anal Mach Intell 42:1194–1204

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Multimedia Tools and Applications
https://doi.org/10.1007/s11042-020-09320-8

A novel intuitionistic fuzzy co-clustering algorithm for brain images

Om Prakash Verma [1] · Heena Hooda [2]

Received: 6 June 2019 / Revised: 10 April 2020 / Accepted: 9 July 2020
Published online: 21 August 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract
Segmentation of Magnetic Resonance Imaging(MRI) Brain images is a very important step in detection of brain tumor. This process is hand operated in labs which is rather an enervating and a long drawn out task and the resultant data so obtained has high degree of erroneity and inconsistency. Hence, automated segmentation systems have become the need of the hour. This paper presents a novel technique to automatically detect brain tumor. Also, segmentation of three main brain tissues is carried out namely white matter, gray matter and cerebrospinal fluid from real time Magnetic Resonance Imaging(MRI) images. The Intuitionistic Fuzzy Set theory is incorporated as it is more suitable for handling uncertainty as compared to fuzzy sets theory. The algorithm is based on the Co-Clustering approach as it offers the advantage of assigning membership functions to both object as well as features. The parameters in the IFCC algorithm are optimized using Particle Swarm Optimization(PSO). The performance of the algorithm is evaluated on the basis of quantitative measures such as match score, accuracy score, dice score and Jaccard's similarity coefficient.

Keywords Magnetic resonance imaging · Intuitionistic fuzzy sets · Fuzzy co-clustering · Tumor detection

1 Introduction

Magnetic Resonance Imaging (MRI) is one of the techniques of imaging technology which is widely used for high quality medical images [6, 16, 22]. Unlike imaging using radiation, in

✉ Heena Hooda
heenahooda@gmail.com

Om Prakash Verma
opverma.dce@gmail.com

[1] GB Pant Govt. Engineering College, Delhi, India

[2] Department Of IT, Delhi Technological University, Delhi, India

which the contrast depends on the different attenuations of the structures being imaged, the contrast in MR images depends on the magnetic properties and number of hydrogen nuclei in the area being imaged. Different contrasts in the area being imaged can be selected by running different sequences with different weightings. Echo Time(TE) is the time between the 90° Radio Frequency pulse and MR signal sampling, corresponding to maximum of echo. The 180° RF pulse is applied at time TE/2. Repetition Time(TR) is the time between 2 excitation pulse. The three main sequences are namely T1-weighted, T2-weighted and Proton Density (PD) weighted depending on the TR and TE sequence. A long TR and a short TE sequence is usually called Proton density -weighted. A short TR and a short TE sequence is usually called T1-weighted. A long TR and a long TE sequence is usually called T2-weighted. MRI is especially used in brain imaging, as it can detect variety of conditions of the brain such as cysts, tumors, bleeding, swelling and discover other such abnormalities of brain. The growth of abnormal cells in the tissues of the brain is known as brain tumor and it is the second leading cause of cancer death. Medical Image Segmentation is primarily used in brain MRI to detect brain tumor, segment brain tissues [7, 8, 31] and partition of brain [24].

Manual Segmentation of MRI images [9] is enervating, long drawn out process, and subjective to the extent that its outcome may vary from individual to individual [30]. Hence, creating an automatic system is the need of the hour to overcome such problems. Although manual segmentation carried out by veterans may give satisfactory results but automated methods can be more promising when dealing with enormous data. To overcome the shortcomings of mapping the tumor manually, the researchers have been keen on developing automatic tumor detection system [26, 34].Automatic tumor detection system involves contouring the tumor part in the image and using various measures for its quantification. In the recent years, various methods have been developed by the researchers for automatic tumor detection but they suffer from some limitations. These techniques are based on morphological operations, edge detection, fuzzy sets theory, classification, region growing, clustering and other image segmentation methods. In 2014, Melegy et al. proposed a new approach by reformulating the popular FCM algorithm to take into account any available information about the class center [15]. Also, local independent projection-based classification (LIPC) method was introduced by Huang et al. [20] to solve the tumor detection problem. The proposed LIPC used local independent projection into the classical classification model, and a novel classification framework was derived. In 2015, Nabizadeh et al. [33] proposed brain tumor detection and segmentation technique for MR images using Gabor wavelet and statistical(GWS) features. Also, Maksoud et al. [1] developed an efficient image segmentation algorithm using KM clustering technique integrated with FCM(KIFCM) algorithm. Sompong et al. [41] proposed brain tumor segmentation method based on gray-level co-occurrence matrix based cellular automata(GLCM-CA) and improved tumor-cut algorithm. A meta-heuristic optimization approach was proposed by Rajinikanth et al. [39] to assist the brain MRI examination. This approach enhances and extracts the tumor core and edema sector from the brain MRI integrating the Teaching Learning Based Optimization (TLBO), entropy value, and level set/active contour based segmentation. In 2018, Deepa et al. [12] proposed adaptive firefly backpropagation neural network(AFBPNN) for detection of brain tumor based on fusion of features using Gaussian radial basis function(GRBF). Javeria et al. [2] also gave a feature fusion based method for finding brain tumor in the year 2018. Various methods have already been proposed for brain tumor detection [19, 27, 38], but they use single membership function for clustering the images which may lead to overlapping clusters, thereby highlighting the need for both feature and object memberships.

Several techniques for MRI brain tissue segmentation have also been proposed and it is an active area of research. Ortiz et al. proposed a hybrid approach based on growing hierarchical self-organising map(GHSOM) classifier and a probability clustering method [37]. Chen et al. modified the objective function of the FCM algorithm to compensate for intensity inhomogeneities by using basis functions and to compensate noise by using improved non-local information [10]. To overcome the problem of FCM algorithm trapping in local minima, in 2013 Benaichouche et al. used a metaheuristic optimization technique for initialisation [5] and incorporated Mahalanobis distance in the objective function to reduce the effect of geometrical shapes of different clusters. Mahmood et al. proposed a framework [29] by integrating Bayesian-based adaptive mean shift, a priori spatial tissue probability maps and FCM. Verma et al. proposed an improved intuitionistic FCM(IIFCM) algorithm [18] for brain image segmentation, which considers the local spatial information in an intuitionistic fuzzy way. The algorithm takes care of the uncertainty with the help of intuitionistic approach and noise effect is taken care of by including local information. Also, Roy et al. performed the three stage repetitive level set(RLS) segmentation method [40] with three membership functions to find out three brain regions from the given image. Convolutional Neural Network(CNN) based method [32] was proposed by Moeskops et al. for automatic segmentation of brain images into tissues. Chen et al. proposed an improved FCM clustering method [11] by using the non-local information and including the covariance and the prior probability in the distance function. Kalaiselvi et al. also modified the FCM algorithm by incorporating spatial parameter for minimizing the objective function of conventional FCM and new weighting parameter for centroid initialization [25]. Dubey et al. [14] gave a novel Rough set based intuitionistic FCM(RIFCM) clustering algorithm for brain tissue classification of MR images. Several bias field based modifications of possibilistic fuzzy C-means algorithms were presented by Aparajeeta et al. [3] for segmenting MRI images. Possibilistic C-means Generalized Rough intutionistic FCM(GRIFCM) algorithm [36] is proposed by Namburu et al. an algorithm to overcome dependency on membership function and parameter tuning. Namburu et al. [35] also proposed algorithm for segmentation of brain MR images using soft set based soft fuzzy rough c-means(SFRCM) clustering.

The segmentation of human brain image from magnetic resonance imaging into three brain tissues: cerebrospinal fluid, gray matter and white matter is one of the important components in computer-aided diagnosis and neuroscience research. It helps to detect different diseases such as tumors, edema, Alzheimer's Disease and Schizophrenia. Due to complicated structure of human brain and absence of well-defined boundary between different tissues, segmentation of brain image is a very difficult task. Hence, working on segmentation of brain images can have good impact in the clinical environment in the real world. There are two major challenges in segmentation of a brain image. First is to handle the uncertainty that arises on boundary between different tissues. Our approach addresses the issue of uncertainty with the help of intuitionistic algorithm. Second is to obtain distinct or non-overlapping clusters which is achieved with the help of co-clustering. We are able to overcome both the challenges with the help of proposed algorithm. In this paper we have proposed a novel Intuitionistic Fuzzy Co-Clustering(IFCC) algorithm for brain tumor detection. The Co-clustering approach offers the advantage of assigning membership functions for both object as well as features. The intuitionistic fuzzy set theory is incorporated as it is more suitable for handling uncertainty as compared to fuzzy sets theory. The required background is discussed briefly in section 2. The proposed clustering algorithm along with the steps involved for tumor detection is explained in section 3. The experimental results are discussed and comparative analysis of the proposed

algorithm is shown in section 4. The paper is concluded in section 5 along with the prospective future work.

2 Background

2.1 Fuzzy sets

Let there be data set $A = \{a_1, a_2, \ldots, a_I\}$where I is the number of elements. Then Zadeh [44] defined a Fuzzy set(FS) B for the set A as in Eq. 1:

$$B = \{\langle a, \mu_B(a)\rangle : \forall a \in A\} \tag{1}$$

where $\mu_B(a)$denotes the membership of element $a \in A$and lies between[0,1]. Also, the non-membership degree is defined as $\nu_B = 1 - \mu_B$. The following properties holds true for two FS B and C [21] are given as:

$$B \subseteq C \; iff \; \forall a \in A, \mu_B(a) \leq \mu_C(a) \tag{2}$$

$$B = C \; iff \; \forall a \in A, \mu_B(a) = \mu_C(a) \tag{3}$$

$$\overline{B} = \{\langle a, 1 - \mu_B(a)\rangle : \forall a \in A\} \tag{4}$$

$$B \cup C = \{\langle a, \max(\mu_B(a), \mu_C(a))\rangle : \forall a \in A\} \tag{5}$$

$$B \cap C = \{\langle a, \min(\mu_B(a), \mu_C(a))\rangle : \forall a \in A\} \tag{6}$$

2.2 Intuitionistic fuzzy sets

In real world when someone needs to take a decision, uncertainty is introduced due to hesitation in expressing their opinion. Intuitionistic Fuzzy sets(IFS) takes into account the degree of hesitation in addition to the membership and non-membership degree as in the FS. The IFS is defined by Atnassov [4] and is given as:

$$I = \{\langle a, \mu_I(a), \nu_I(a)\rangle : \forall a \in A\} \tag{7}$$

with constraint

$$0 \leq \mu_I(a) + \nu_I(a) \leq 1 \tag{8}$$

where $\mu_I(a)$ and $\nu_I(a)$denotes the membership and non-membership degree of element $a \in A$and lies between[0,1]. The hesitation degree of the element $a \in A$ is given by eq. 9.

$$\pi_I(a) = 1 - \mu_I(a) - \nu_I(a) \tag{9}$$

If the value of $\pi_I(x) = 0$in eq. 9, then the IFS is reduced to FS.

The properties [43] that holds true for two IFS H and I are given as:

$$I \subseteq H \text{ iff } \forall a \in A, \mu_I(a) \leq \mu_H(a) \text{and} \nu_I(a) \geq \nu_H(a) \tag{10}$$

$$I = H \text{ if f } \forall a \in A, \mu_I(a) = \mu_H(a) \text{and} \nu_I(a) = \nu_H(a) \tag{11}$$

$$\overline{I} = \{\langle a, \nu_I(a), \mu_I(a)\rangle : \forall a \in A\} \tag{12}$$

$$I \cup H = \{\langle a, \max(\mu_I(a), \mu_H(a)), \min(\nu_I(a), \nu_H(a))\rangle : \forall a \in A\} \tag{13}$$

$$I \cap H = \{\langle a, \min(\mu_I(a), \mu_H(a)), \max(\nu_I(a), \nu_H(a))\rangle : \forall a \in A\} \tag{14}$$

For the construction of IFS, the negation function is given by Sugeno [42] as:

$$\theta_\lambda(a) = \frac{1-a}{1+\lambda a} \qquad \lambda > 0 \tag{15}$$

The IFS I is defined using the Sugeno's negation function as in eq. 16.

$$I = \left\{ \left\langle a, \mu_I(a), \nu_I(a) = \frac{1-\mu_I(a)}{1+\lambda(\mu_I(a))} \right\rangle : \forall a \in A \right\} \tag{16}$$

2.3 Intuitionistic Fuzzification of images

The intuitionistic fuzzification of images gives the appropriate membership and non-membership functions to describe the properties of the image in terms of IFS. A digital image I in the form of IFS can be represented as

$$I = \left\{ \left\langle a_j, \mu_I(a_j), \nu_I(a_j) \right\rangle : j = 1, 2, \ldots, N \right\} \tag{17}$$

where N is the total number of pixels of the image (N=N1*N2 is the size of the image), with L number of gray levels varying from 0 to L-1,the membership degree of pixels a_j is indicated by $\mu_I(a_j)$ and non-membership degree of pixels a_j is indicated by $\nu_I(a_j)$ .After normalizing the intensity level we obtain the membership degree of every pixel $x_j$of the image as given in eq. 18.

$$\mu_I(a_j) = \frac{a_j - (a_j)_{\min}}{(a_j)_{\max} - (a_j)_{\min}}, \quad j = 1, 2, \ldots, N, \quad x_j \in \{0, \ldots, L-1\} \tag{18}$$

where $(a_j)_{\min}$ and $(a_j)_{\max}$ are the minimum and maximum values of the gray levels of the image respectively.

2.4 Fuzzy co-clustering for images algorithm

The aim of Co-clustering is to group both objects and features concurrently. Both the object and feature memberships in the proposed method are fuzzy,i.e. the object membership is calculated

when different clusters compete for a datapoint and feature memberships are defined when different features compete for a cluster. Thus we have two constraints on the two fuzzy memberships(object and feature memberships) in our method. We have used FCCI algorithm [17] instead of FCM in our work as it offers several advantages over the latter. First advantage is that it performs well in high dimensions and it also gives us better clusters as compared to FCM clustering.

The objective function of the FCCI algorithm is given in eq. 19. We can observe that there are three terms in the equation, the first term is $\sum_{i=1}^{C}\sum_{j=1}^{N}\sum_{k=1}^{R} u_{ij}f_{ik}D_{ijk}$ which accounts for dissimilarity between the data point and cluster centre of the feature matrix by incorporating the distance function. In the next two terms regularization of objective function is done using entropies. The entropy of objects is taken as $\sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}logu_{ij}$ and entropy of features is taken as $\sum_{i=1}^{C}\sum_{k=1}^{R} f_{ik}logf_{ik}$. The images captured for applying the FCCI algorithm are first converted from RGB to CIELAB color space and two features A and B of the CIELAB color space are then further used to perform the clustering. The luminance vector obtained after transformation is not taken into consideration in the clustering process as it is only used to determine the darkness or fairness of the image segments and does not have any effect on the segmentation process.

The objective function J given in eq. 19 is minimized in the FCCI algorithm [17]:

$$J = \sum_{i=1}^{C}\sum_{j=1}^{N}\sum_{k=1}^{R} u_{ij}f_{ik}D_{ijk} + T_U\sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}logu_{ij} + \quad T_V\sum_{i=1}^{C}\sum_{k=1}^{R} f_{ik}logf_{ik} \tag{19}$$

The total membership of all attributes in a cluster must sum to 1, hence eq. 19 is subjected to the following constraints:

$$\sum_{i=1}^{C} u_{ij} = 1, \qquad u_{ij}\in[0,1], \quad \forall j = 1,\ldots,N \tag{20}$$

$$\sum_{k=1}^{R} f_{ik} = 1, \qquad f_{ik}\in[0,1], \quad \forall i = 1,\ldots,C \tag{21}$$

where, C is the number of clusters, N is the number of data points i.e., N = N1 × N2 (N1*N2 is the dimension of the input image), R is the number of color features, u_{ij} is membership function of data point j to cluster $i$given in eq. 22, f_{ik} is membership function of feature k to cluster $i$given in eq. 23, T_U and T_V are the weighting parameters that determine the degree of fuzziness and D_{ijk} is the Euclidean distance between data point x_{jk} and cluster centroid p_{ik} given in eq. 24.

$$u_{ij} = \frac{e^{\left(-\sum_{k=1}^{R}\frac{f_{ik}D_{ijk}}{T_U}\right)}}{\sum_{i=1}^{C} e^{\left(-\sum_{k=1}^{R}\frac{f_{ik}D_{ijk}}{T_U}\right)}} \tag{22}$$

$$f_{ik} = \frac{e^{\left(-\sum_{j=1}^{N}\frac{u_{ij}D_{ijk}}{T_V}\right)}}{\sum_{k=1}^{R} e^{\left(-\sum_{j=1}^{N}\frac{u_{ij}D_{ijk}}{T_V}\right)}} \tag{23}$$

$$D_{ijk} = d^2\left(x_{jk}, p_{ik}\right) = \left(x_{jk} - p_{ik}\right)^2 \tag{24}$$

where,

$$p_{ik} = \frac{\sum_{j=1}^{N} u_{ij} x_{jk}}{\sum_{j=1}^{N} u_{ij}} \tag{25}$$

2.5 Intuitionistic Fuzzification of images for co-clustering

The membership, non-membership and hesitation degree of every pixel $x_{jk}$of the matrix for k features in terms of normalize intensity level are illustrated in eqs. 26–28

$$\mu_I\left(x_{jk}\right) = \frac{x_{jk} - \left(x_{jk}\right)_{\min}}{\left(x_{jk}\right)_{\max} - \left(x_{jk}\right)_{\min}}, \quad j = 1, 2, \ldots, N, k = 1, 2 \;\; x_j \in \{0, \ldots, L-1\} \tag{26}$$

$$\nu_I\left(x_{jk}\right) = \frac{1 - \mu_I\left(x_{jk}\right)}{1 + \lambda\left(\mu_I\left(x_{jk}\right)\right)} \tag{27}$$

$$\pi_I\left(x_{jk}\right) = 1 - \mu_I\left(x_{jk}\right) - \nu_I\left(x_{jk}\right) \tag{28}$$

where $(x_{jk})_{\min}$ and $(x_{jk})_{\max}$ are the minimum and maximum values of the gray levels of the image respectively.

3 Proposed work

3.1 Intuitionistic fuzzy co-clustering algorithm

Major challenge in segmentation of brain image is to handle the uncertainty that arises from boundary between different tissues. To handle this type of uncertainty, Atanassov suggested the generalization of fuzzy sets theory, known as intuitionistic fuzzy sets (IFS). Distinct or non-overlapping clusters are achieved with the help of Fuzzy co-clustering. We are able to overcome both the challenges with the help of proposed algorithm.

Let $X^{IFS} = \left\{x_1^{IFS}, x_2^{IFS}, \ldots, x_N^{IFS}\right\}$ be the set of N data points of the image I(total number of pixels N=N1*N2, where N1 and N2 are rows and columns of the image respectively) and R be the number of features linked with each data point. Then, to denote each data point for a particular feature we can write k^{th} feature of the jth data point as, x_{jk}.

Let $(P^{IFS})^{(b)} = ((p_{1k})^{(b)}, (p_{2k})^{(b)}, \ldots, (p_{ik})^{(b)})$ be the set of feature centroid for each feature k of cluster i and $D_{ijk}^{IFS} = Dist\left(x_{jk}^{IFS}, p_{ik}^{IFS}\right)$ be the square of the Euclidean distance between feature data point $x_{jk}$and feature cluster centroid p_{ik}.

The distance is given by the equation

$$D_{ijk}^{IFS} = d^2\left(x_{jk}^{IFS}, p_{ik}^{IFS}\right) = \left(x_{jk}^{IFS} - p_{ik}^{IFS}\right)^2 \tag{29}$$

Let object membership function matrix is $U = \{u_{ij}\}$ where u_{ij} is the object membership of the data point j to cluster i and the feature membership matrix is $F = \{f_{ik}\}$ where f_{ik} is the feature membership of the feature k to the cluster i.

The minimization optimization problem of the IFCC algorithm is defined as:

$$J_{IFCC}\left(U, F, P^{IFS} : X^{IFS}\right) = \sum_{i=1}^{C}\sum_{j=1}^{N}\sum_{k=1}^{R} u_{ij} f_{ik} D\left(x_{jk}^{IFS}, p_{ik}^{IFS}\right) + T_U \sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij} \log u_{ij} + T_f \sum_{i=1}^{C}\sum_{k=1}^{R} f_{ik} \log f_{ik} \tag{30}$$

subject to the constraints:

$$\sum_{i=1}^{C} u_{ij} = 1. \qquad u_{ij} \in [0, 1], \forall j = 1, \ldots, N \tag{31}$$

$$\sum_{k=1}^{R} f_{ik} = 1. \qquad f_{ik} \in [0, 1], \forall i = 1, \ldots, C \tag{32}$$

Now, we use Lagrangian multiplier method to solve the optimization problem and constrained optimization problem can now be formed as:

$$J_{IFCC}^{*}\left(U, F, P^{IFS} : X^{IFS}\right) = \sum_{i=1}^{C}\sum_{j=1}^{N}\sum_{k=1}^{R} u_{ij} f_{ik} D\left(x_{jk}^{IFS}, p_{ik}^{IFS}\right) + T_U \sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij} \log u_{ij} + \mathrm{T_f} \sum_{i=1}^{C}\sum_{k=1}^{R} f_{ik} \log f_{ik} + \sum_{j=1}^{N} \lambda_j \left(\sum_{i=1}^{C} u_{ij} - 1\right) + \sum_{i=1}^{C} \gamma_i \left(\sum_{k=1}^{R} f_{ik} - 1\right) \tag{33}$$

where $\lambda_j$and $\gamma_i$are multiplier constants.

Now, we calculate the partial derivatives of $J^{*}_{IFCC}(U, F, P^{IFS} : X^{IFS})$with respect to U and setting the gradient to zero we get

$$u_{ij} = \frac{e^{\left(-\sum_{k=1}^{R} \frac{f_{ik} D\left(x_{jk}^{IFS}, p_{ik}^{IFS}\right)}{T_U}\right)}}{\sum_{i=1}^{C} e^{\left(-\sum_{k=1}^{R} \frac{f_{ik} D\left(x_{jk}^{IFS}, p_{ik}^{IFS}\right)}{T_U}\right)}} \tag{34}$$

Similarly, we calculate the partial derivatives of $J^{*}_{IFCC}(U, F, P^{IFS} : X^{IFS})$with respect to F and setting the gradient to zero we get

$$f_{ik} = \frac{e^{\left(-\sum_{j=1}^{N} \frac{u_{ij} D\left(x_{jk}^{IFS}, p_{ik}^{IFS}\right)}{T_V}\right)}}{\sum_{k=1}^{R} e^{\left(-\sum_{j=1}^{N} \frac{u_{ij} D\left(x_{jk}^{IFS}, p_{ik}^{IFS}\right)}{T_V}\right)}} \tag{35}$$

Now, taking the partial derivatives of $J^*_{IFCC}(U, F, P^{IFS} : X^{IFS})$with respect to $\mu(p_{ik})$, $\nu(p_{ik})$ and $\pi(p_{ik})$and setting the gradient to zero we get

$$\mu(p_{ik}) = \frac{\sum_{j=1}^{N} u_{ij}\mu(x_{jk})}{\sum_{j=1}^{N} u_{ij}} \tag{36}$$

$$\nu(p_{ik}) = \frac{\sum_{j=1}^{N} u_{ij}\nu(x_{jk})}{\sum_{j=1}^{N} u_{ij}} \tag{37}$$

$$\pi(p_{ik}) = \frac{\sum_{j=1}^{N} u_{ij}\pi(x_{jk})}{\sum_{j=1}^{N} u_{ij}} \tag{38}$$

The centroids $P^{IFS}$of the IFCC algorithm are given as:

$$p_{ik}^{IFS} = (\mu(p_{ik}), \nu(p_{ik}), \pi(p_{ik})) \tag{39}$$

The IFCC algorithm is summarized as:

Step 1: Fix the number of cluster $C = 2$, number of features $k = 2$, fuzzifier constant $m = 1$ and termination criterion $\varepsilon=10^{-2}$.

Step 2: Calculate $x_{jk}^{IFS} = \left(\mu(x_{jk}), \nu(x_{jk}), \pi(x_{jk})\right)$, for $j = 1,2,...,N$ and $k = 1,2$ using eqs. 26–28.

Step 3: Initialize the membership matrix $U^{(o)}$ with setting the loop counter $b = 0$.

Step 4: Update the centroids,where $(P^{IFS})^{(b)} = ((p_{1k})^{(b)}, (p_{2k})^{(b)}, \ldots, (p_{ik})^{(b)})$, where $\left(p_{ik}^{IFS}\right)^{(b)} = \left((\mu(p_{ik}))^{(b)}, (\nu(p_{ik}))^{(b)}, (\pi(p_{ik}))^{(b)}\right)$ using eqs. 36–38.

Step 5: CalculateD_{ijk} and f_{ik} using eqs. 29 and 35.

Step 6: Update the membership matrix $U^{(b)} = ((u_{ij})_{C\times N})^b$ using eq. 34.

Step 7: If $\max\|U^{(b+1)} - U(b)\| < \varepsilon$then stop, otherwise set counter b = b + 1 and go to step 4.

3.2 Tumor detection using IFCC

The proposed tumor detection system consists of three modules namely pre-processing, building the intracranial mask and tumor extraction. In the first module, we process the brain images before giving them as input to the system. Then, the intracranial mask is build so as to remove undesirable part from MRI images in the next module. And finally the proposed algorithm is applied to extract the tumor using the intracranial mask.

3.2.1 Pre-processing

In the first step the MRI scan images are fed to the median filter to remove the effect of various types of noises such as irregularities. The main objective of performing this pre-processing step is to guarantee the determination of the tumor accurately.

3.2.2 Building the intracranial mask

The intent of this step in the proposed system is to derive the intra-cranial mask from the entire image. The extra-cranial region comprises of bones which do not contain tumor and hence its inclusion in the brain image is insignificant in segmentation of tumor. After studying and analyzing the images we see that there is remarkable difference between the intensity of intra-cranial and extra-cranial region. This difference is taken as a measure to derive the intracranial mask and remove the extra cranial mask from the MRI images. A threshold value is used to isolate the intra-cranial and extra-cranial masks. The threshold value at which the intra-cranial and extra-cranial regions are separated is 0.5. The extra-cranial mask obtained is subtracted from the original image to obtain the desired intra-cranial mask as shown in figure 1. If *G(a,b)* be the original MRI image and *E(a,b)* is the extracranial mask obtained after thresholding then, intracranial mask *F(a,b)* is determined using eq. 40:

$$G(a,b) = F(a,b) - E(a,b) \tag{40}$$

where, a is the number of rows and b is the number of columns in the image.

The algorithm for obtaining the desired intracranial mask is described below:

Step 1: Read the input MRI image as shown in Fig. 1(a).
Step 2: Apply thresholding on the input image with threshold value t = 0.5. As a result, we get two images as shown in Fig. 1(b) and (c).
Step 3: From the images obtained in step 2 choose the image with higher intensity values. Hence, we choose Fig. 1(b)
Step 4: From the chosen image, i.e. Figure 1(b) take only the connected component region i.e. the extracranial mask as shown in Fig. 1(d).
Step 5: The extracranial mask obtained in Fig. 1(d) is subtracted from the original image to obtain the desired intracranial mask as shown in Fig. 1(e).

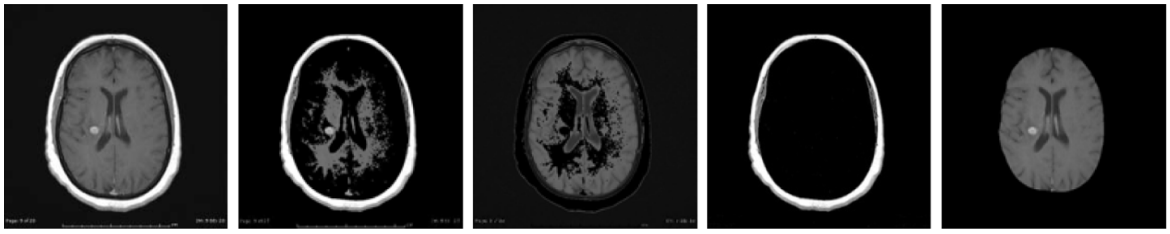
3.2.3 Initialization of IFCC using particle swarm optimization

PSO is a population-based stochastic optimization algorithm proposed for the first time by Kennedy and Eberhart [28], inspired by bird flocking and fish schooling. The problem is tackled by considering a population(particles), where each particle is a potential solution to the problem. Initial positions and velocities of the particles are chosen randomly. In the commonly used standard PSO, each particle's position is updated at each iteration step according to its own personal best position and the best solution of the swarm. The evolution of the swarm is governed by the following equations:

$$\begin{aligned} VEL(iter+1) = w.VEL(iter) + c1.rand1.(pbest(iter) - POS(iter)) \\ + c2.rand2.(gbest(iter) - POS(iter)) \end{aligned} \tag{41}$$

$$POS(iter+1) = POS(iter) + VEL(iter+1) \tag{42}$$

where, *POS* is the position of the particle, *VEL* is the velocity of the particle, w is the inertia weight, *pbest* is the personal best position of the particle, *gbest* is the global best position of the



(a) (b) (c) (d) (e)

Fig. 1 (a) Original MRI image, (b-c) Images after thresholding, (d) Extracranial mask, (e) Intracranial Mask

swarm, *rand*1, *rand*2 are random values between 0 and 1, $c1$, $c2$ are positive constants which determine the impact of the personal best solution and the global best solution on the search process, respectively, and *iter* is the iteration number.

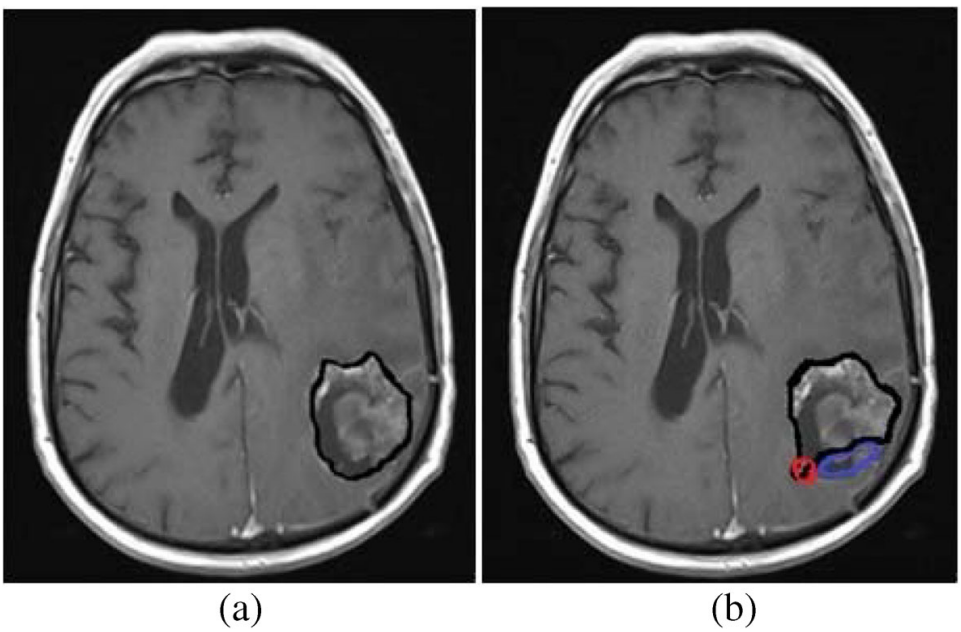
The parameters T_U and T_V in the IFCC algorithm are optimized for each image using PSO algorithm [28]. The parameters are initialized with the values $T_U = 10$ and $T_V = 10^7$ as determined experimentally. The choice of these initial values is made by conducting a set of random experiments by hit and trial. The number of particles are initialized as 10 and maximum number of iterations as 100 since all our experiments are found to converge within 100 iterations.

3.2.4 Detection of brain tumor

Given an intracranial mask, our sole aim is now to detect the tumor. IFCC algorithm is applied for the clustering of the given MRI images. The tumor can be detected by using the image obtained after applying the clustering technique and the intra-cranial mask.

The algorithm of the proposed approach is described in the steps below:

Step 1: Read the input color MRI scan image.
Step 2: Convert the input color image into grayscale image.
Step 3: Apply median filter to remove noise.
Step 4: The desired intracranial mask is obtained as shown in Fig. 1.



(a) (b)

Fig. 2 **a** Ground truth tumor **b** True positive, false positive and false negative markings

Table 1 Quantitative results of tumor detection using IFCC algorithm

S.No.	Patient No.	Ground Truth	True Positive	False Negative	False Positive	Match Score	Accuracy Score	Dice Coefficient	Jaccard Score
1	1	1928	1800	128	0	0.93	0.90	0.97	0.93
2	1	4360	4194	166	0	0.96	0.94	0.98	0.96
3	1	4916	4867	49	0	0.99	0.99	0.99	0.99
4	1	2056	2056	0	0	1.00	1.00	1.00	1.00
5	1	1150	1150	0	0	1.00	1.00	1.00	1.00
6	2	1651	1420	231	48	0.86	0.78	0.92	0.86
7	2	1884	1866	18	0	0.99	0.99	1.00	0.99
8	2	2324	2312	12	157	0.99	0.95	1.00	0.99
9	2	514	514	0	0	1.00	1.00	1.00	1.00
10	3	8265	5951	2314	102	0.72	0.57	0.84	0.72
11	3	10,102	7880	2222	481	0.78	0.65	0.88	0.78
12	4	744	707	37	0	0.95	0.93	0.97	0.95
13	4	1498	1498	0	0	1.00	1.00	1.00	1.00
14	5	5212	4706	506	521	0.90	0.80	0.95	0.90
15	5	5680	5413	267	0	0.95	0.93	0.98	0.95
16	5	4794	4794	0	0	1.00	1.00	1.00	1.00
17	6	13,422	13,422	0	0	1.00	1.00	1.00	1.00
18	6	11,022	10,912	110	3306	0.99	0.84	0.99	0.99
19	6	9164	9073	91	2749	0.99	0.84	1.00	0.99
20	7	1983	1884	99	0	0.95	0.93	0.97	0.95
21	7	1802	1442	360	0	0.80	0.70	0.89	0.80
22	7	1686	1686	0	33	1.00	0.99	1.00	1.00
23	7	2704	2569	135	0	0.95	0.93	0.97	0.95
24	7	3845	3845	0	0	1.00	1.00	1.00	1.00
25	7	2705	2705	0	1082	1.00	0.80	1.00	1.00
26	8	5500	4510	990	0	0.82	0.73	0.90	0.82
27	8	6705	5834	871	335	0.87	0.78	0.93	0.87
28	8	6171	4937	1234	0	0.80	0.70	0.89	0.80
29	8	7412	5930	1482	1482	0.80	0.60	0.89	0.80
30	9	1315	1302	13	0	0.99	0.99	1.00	0.99
31	9	1968	1575	393	0	0.80	0.70	0.89	0.80
32	10	6748	6681	67	0	0.99	0.99	1.00	0.99
33	10	5758	5471	287	0	0.95	0.93	0.97	0.95
34	10	3495	3461	34	0	0.99	0.99	1.00	0.99
35	11	4295	4081	214	0	0.95	0.93	0.97	0.95
36	11	4434	4434	0	0	1.00	1.00	1.00	1.00
37	11	4360	4360	0	0	1.00	1.00	1.00	1.00
38	11	2903	2323	580	0	0.80	0.70	0.89	0.80
39	12	2362	2339	23	0	0.99	0.99	1.00	0.99
40	13	1889	1871	18	0	0.99	0.99	1.00	0.99
41	13	2891	2313	578	0	0.80	0.70	0.89	0.80
42	14	6288	5031	1257	0	0.80	0.70	0.89	0.80
43	15	5028	4719	309	26	0.93	0.90	0.97	0.94
44	15	4824	4342	482	0	0.90	0.85	0.95	0.90
45	15	2918	2918	0	0	1.00	1.00	1.00	1.00
46	16	1496	1378	118	0	0.92	0.88	0.96	0.92
47	16	4913	4581	332	0	0.93	0.89	0.97	0.93
48	16	5295	5295	0	52	1.00	0.99	1.00	1.00
49	16	4106	3833	273	0	0.93	0.90	0.97	0.93
50	16	1754	1611	143	38	0.92	0.87	0.96	0.92

Step 5: Apply IFCC algorithm explained in section 3.1 with values of $C = 2$ to the intracranial mask obtained in step 4.

Table 2 DC values of proposed technique and other segmentation algorithms

S.No.	Ground Truth	FCM		IFCM		IIFCM		IFCC	
		True Positive	Dice Score	True Positive	Dice Score	True Positive	Dice Score	True Positive	Dice Score
1	1928	2376	0.9	1786	0.96	1600	0.91	1800	0.97
2	4360	4867	0.95	5990	0.84	4142	0.97	4194	0.98
3	2324	1866	0.89	1765	0.86	2301	1	2312	1
4	514	683	0.86	528	0.99	514	1	514	1
5	5212	4170	0.89	4023	0.87	4278	0.9	4706	0.95
6	5680	5031	0.94	3805	0.8	5112	0.95	5413	0.98
7	4794	4983	0.98	4126	0.93	4794	1	4794	1
8	5028	4676	0.96	4342	0.93	4526	0.95	4719	0.97
9	4913	5231	0.97	4451	0.95	4451	0.95	4581	0.97
10	5295	3833	0.84	5371	0.99	5182	0.99	5295	1

Step 6: Select the clustered image which has more number of pixels with values close to 1 as that cluster will contain tumor.

Step 7: Apply morphological operation close to remove unwanted part and to get the exact shape of the tumor.

Step 8: Calculate the area of tumor by calculating the number of white pixels in the tumorous image.

4 Results & discussion

The algorithm is implemented on personal computer (2.0 GHz CPU, 2GB RAM) using MATLAB 7.9.0 (2009b). The tumorous images have been provided from a database of T1 weighted brain MRI images of 20 patients by the radiologist of Rajiv Gandhi Cancer Institute & Research Centre(RGCI&RC), Delhi, India. Each patient dataset had 20 MRI images in form of horizontal slices of brain resulting in a total of 400 images. The proposed scheme is applied on 50 tumorous MRI images obtained from real time database of RGCI&RC. Also, the real time database of Brain MRI images has been taken from Insight Journal. Insight Journal is an Open Access on-line publication that covers domain of medical image processing and visualization. One of its journals is MIDAS. Midas community include National Alliance for medical image computing (NAMIC) which presents the data for two autistic and two normal children (male and female) (http://www.insight-journal.org/midas/community/). The data is re-collected after a time span of 2 years. Three types of MRI scanning is presented i.e. T1 weighted, T2 weighted, PD weighted images. Coronal slices are obtained with slice thickness of 1.5mms. Also the tissue segmentation label map is presented with the database. This tissue label is atlas based segmentation by making use of expectation-maximization scheme. After mapping the tumor on the original MRI image using the proposed tumor detection system, the calculated tumor area is compared with the ground truth tumor manually mapped by the radiologist. The radiologist analyzed the outputs based on four terminologies; ground truth, true positive, false positive and false negative as shown in Fig. 2.

The area accurately measured is termed as *True Positive* and the errors encountered are termed as *False Positive* and *False Negative*. False positive counts the total number of pixels that our system classifies but is not present in ground truth and True negative counts the total

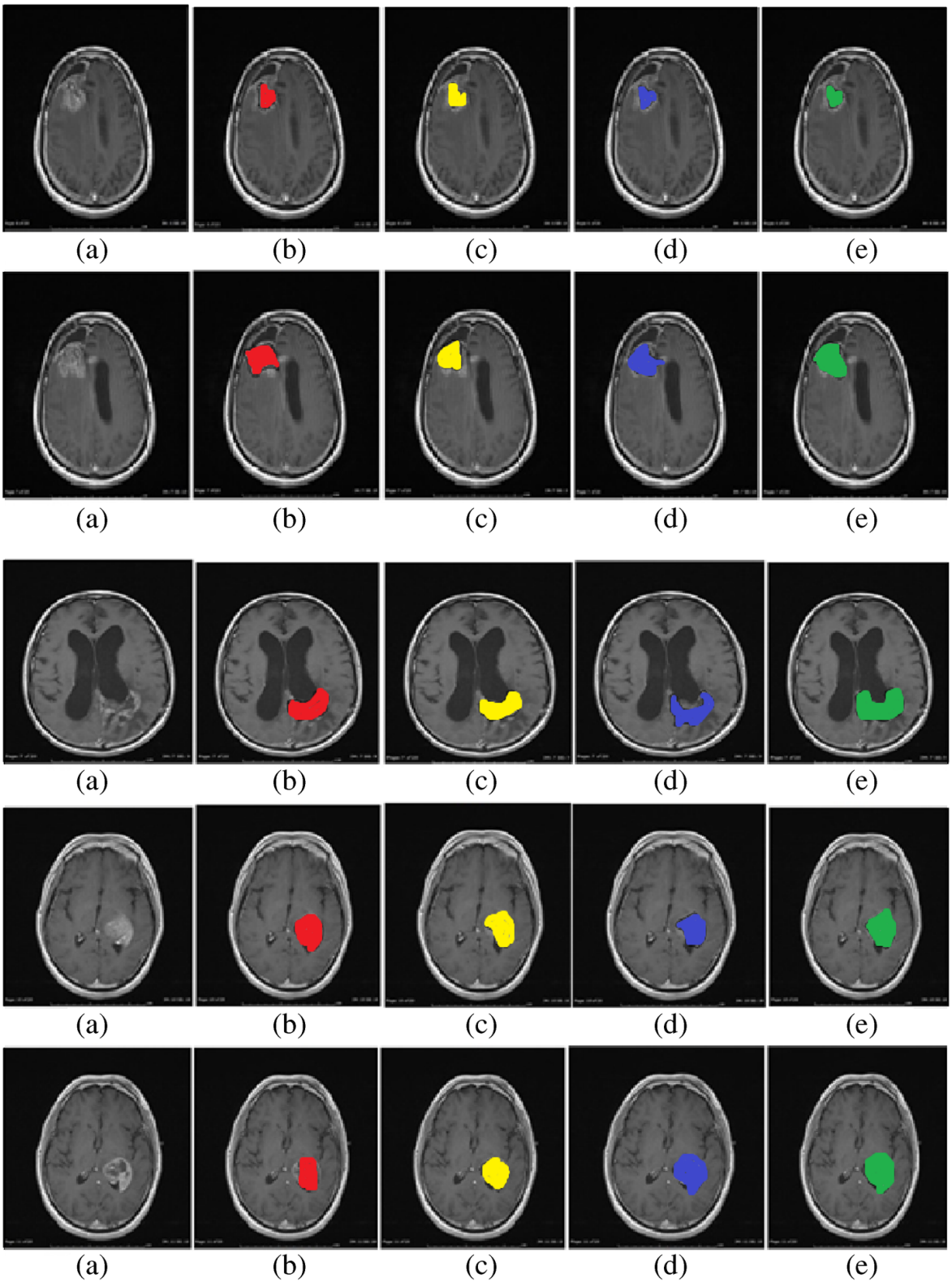


Fig. 3 (a) Input MRI image, Tumor mapping with (b) FCM (c) IFCM (d) IIFCM (e) IFCC

number of pixels that our system doesn't classifies but is present in the ground truth as belonging to a particular cluster. To evaluate how well the system performs in segmenting the tumor, we use four quantitative measures namely, Match Score, Accuracy Score, Dice Coefficient(DC) and Jaccard's similarity(JS) score which are formulated from the parameters given by the radiologist and can be seen in Eqs. 43 to 46. The number of True Positive pixels

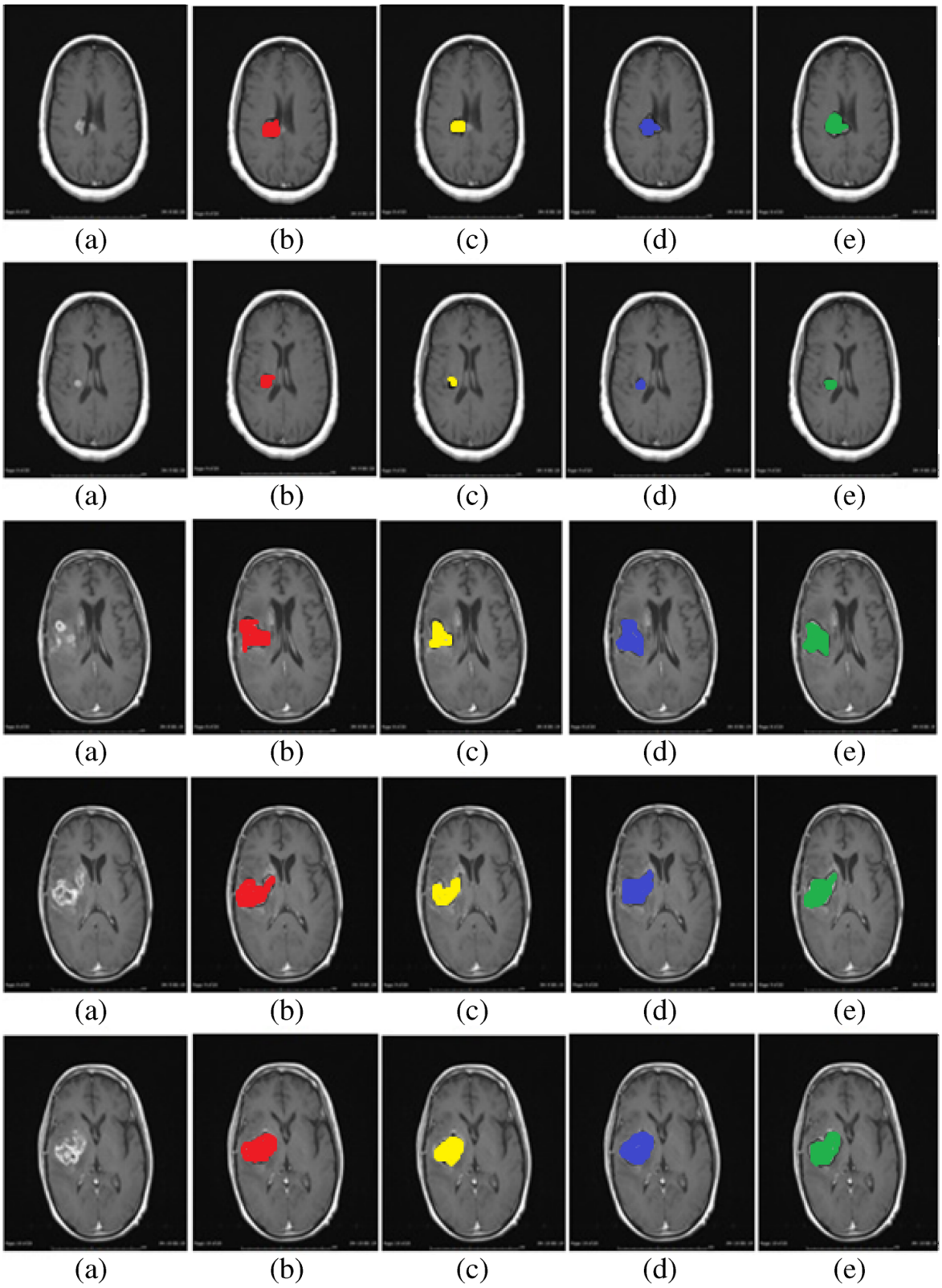


Fig. 3 (continued)

divided by the number of ground truth pixels results in *Match Score* which indicates the performance of proposed algorithm in comparison with the ground truth given by the radiologist. *Accuracy Score* is calculated to account for the presence of error in segmenting tumor and therefore the formula incorporates both false positive and false negative.

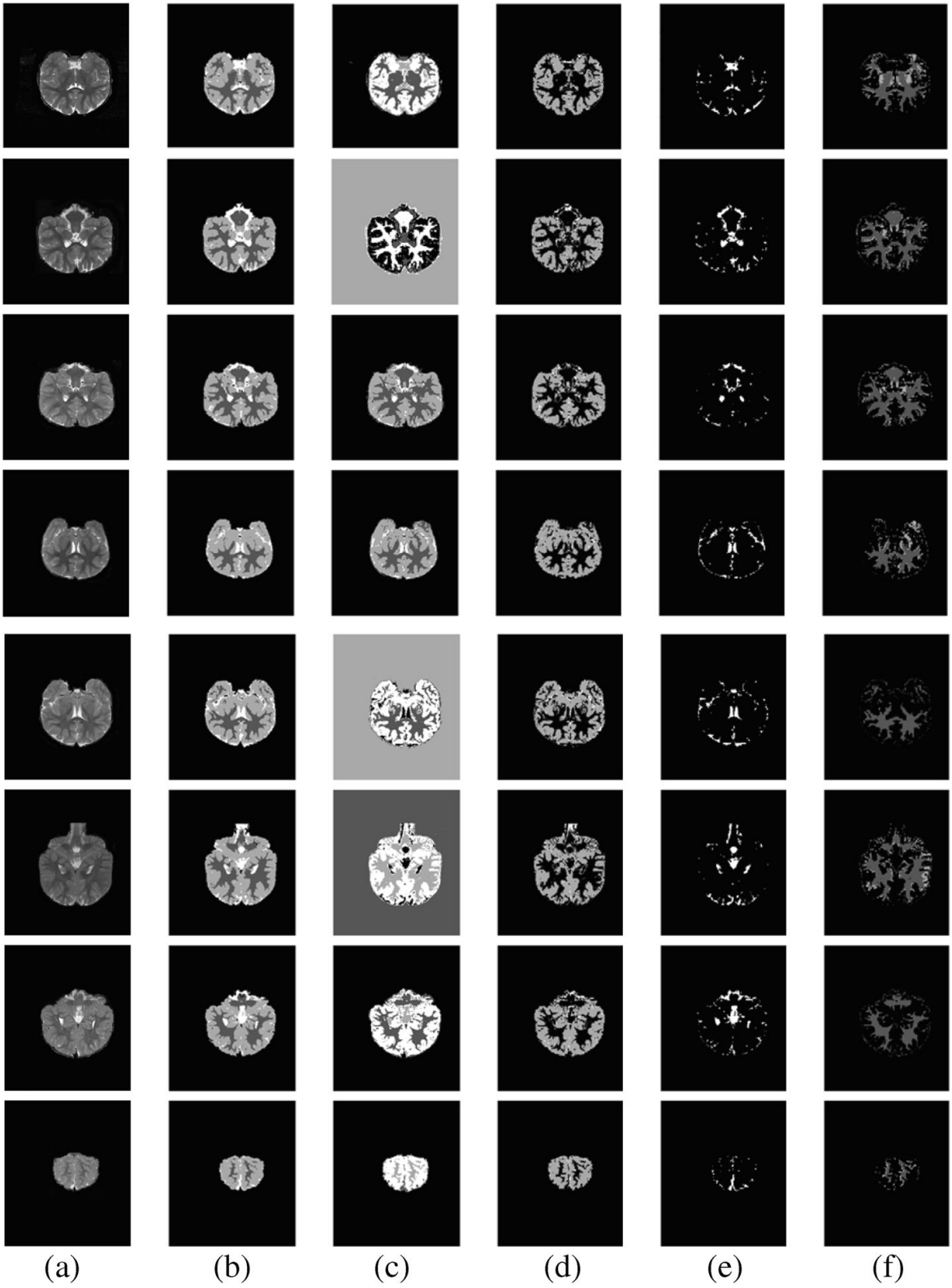


Fig. 4 **a** MRI slices, **b** Ground truth, **c** Segmentation result, **d** Gray matter, **e** CSF, **f** White matter

$$\text{Match Score} = \text{True positive}/\text{Ground truth} \tag{43}$$

$$\text{Accuracy Score} = \frac{\text{True positive}–0.5(\text{False positive} + \text{False negative})}{\text{Ground truth}} \tag{44}$$

The DC [13] is defined as

$$DC(S, W) = 2(|S \cap \mathrm{W}|)/(|S| + |W|) \tag{45}$$

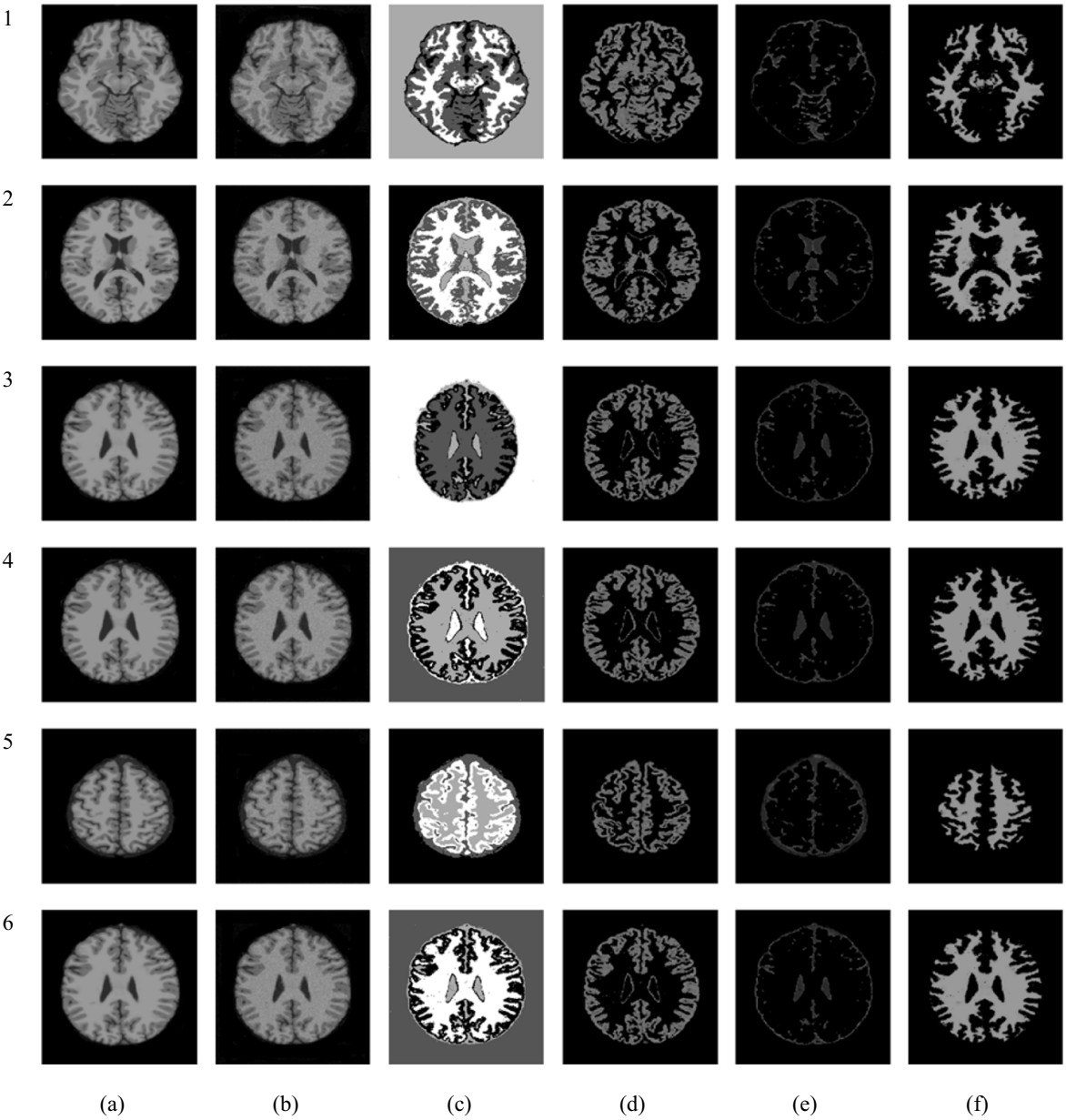


Fig. 5 (**a**) Normal Brain, (**b**) Normal brain (noise 3% for 1 to 3) and (noise 5% for 4 to 6), (**c**) Segmented result, (**d**) Gray matter, (**e**) CSF, (**f**) White matter

Table 3 Comparative analysis of MRI image segmentation algorithms

Segmentation Technique	Dice Score	Jaccard's Score	Error Percentage(%)
PIGFCM [15]	0.91	0.85	8.1
LIPC [20]	0.83	0.75	16.4
GWS [33]	0.80	0.74	17
KIFCM [1]	0.88	0.82	9.2
GLCM-CA [41]	0.92	0.86	7.4
TLBO [39]	0.91	0.89	6.0
AFBPNN [12]	0.92	0.86	8.0
GRBF [2]	0.93	0.88	7.5
Proposed Algorithm: IFCC	0.96	0.92	3.7

Table 4 Brain Tissue Segmentation results of IFCC algorithm

Patient #	Ground Truth	True positive	False Negative	False Positive	Match Score	Accuracy Score	Dice Score	Jaccard Score
P11	34,482	29,779	4703	2172	0.8636	0.7639	0.9268	0.8636
P12	33,936	29,829	4017	4449	0.8790	0.7542	0.9356	0.8790
P21	33,056	30,733	2323	3581	0.9297	0.8404	0.9636	0.9297
P22	28,503	24,764	3739	3946	0.8688	0.7340	0.9298	0.8688
P31	33,528	28,973	4555	5197	0.8641	0.7187	0.9271	0.8641
P32	39,125	34,822	4303	4442	0.8900	0.7783	0.9418	0.8900
P41	30,248	28,660	1588	3036	0.9475	0.8711	0.9730	0.9475
P42	10,960	10,183	777	809	0.9291	0.8568	0.9633	0.9291

where S and W are two sets. The value of DC lies between 0 and 1. 0 signifies no match/overlap and 1 signifies complete match/overlap. In terms of segmentation, *'S'* represents the segmented image containing tumor and '*W*' represents the ground truth given by the radiologist. The JS, also termed as Jaccard index the coefficient was originally given by Paul Jaccard [23], is a statistic used for comparing the similarity and diversity of sample sets. The JS measures similarity between finite sample sets, and is defined as the size of the intersection divided by the size of the union of the sample sets as given in eq. 46:

$$JS = (|S \cap W|)/(|S \cup W|) \tag{46}$$

The value of JS also lies within 0 and 1. 0 signifies no match/overlap and 1 signifies complete match/overlap. The coefficient values of 1 are desired for dice score as well as JS. Relationship between DC and JS is as follows:

$$DC = 2*JS/(1 + JS) \tag{47}$$

The match score, accuracy score, DC and JS have been calculated for 50 images using IFCC algorithm and shown in Table 1. The more the value of DC is close to 1, the better is the segmentation accuracy. It is studied that dice coefficient's value >0.7 implies good segmentation. Similar to DC is the JS measure. Both these parameters are not sensitive to volume overestimations and underestimations. However, DC is more famous than JS as it may sometimes result in a mismatch when there is a strong volumetric overlap. We have presented the estimated results using JS as well to have a better outlook of classification system. As we can see from the table that the proposed system has an average match score of 0.92 which indicates high degree of similarity between the ground truth and the brain tumor segmented using the proposed algorithm. The

Table 5 DC values for three segmentation of brain tissues for Fig. 4 using IFCC approach

Patient #	GM	WM	CSF	Avg.
P11	0.991	0.969	0.89	0.950
P12	0.972	0.967	0.881	0.940
P21	0.969	0.975	0.928	0.957
P22	0.99	0.941	0.881	0.937
P31	0.984	0.959	0.892	0.945
P32	0.996	0.972	0.873	0.947
P41	0.989	0.952	0.918	0.953
P42	0.992	0.973	0.937	0.967
Average	**0.985**	**0.964**	**0.900**	**0.950**

Table 6 DC values for segmentation of brain tissues for Fig. 5 using IFCC for Brain Web database

Slice#	Noise	GM	WM	CSF	Avg.
60	3%	0.991	0.911	0.812	0.905
88	3%	0.962	0.963	0.89	0.938
100	3%	0.947	0.97	0.837	0.918
AVG.		**0.97**	**0.948**	**0.846**	**0.920**
99	5%	0.978	0.967	0.841	0.929
126	5%	0.953	0.902	0.874	0.910
100	5%	0.941	0.962	0.835	0.913
AVG.		**0.957**	**0.943**	**0.85**	**0.917**

average accuracy score is 0.87, as it also incorporates the errors that occurred in extracting tumor. The average DC and JS have been reported 0.96 and 0.92 respectively.

It can be observed from Table 2 that the values of dice and Jaccard's Score are highest for the proposed algorithm, hence IFCC outperforms other segmentation algorithms for detection of brain tumor from MRI images. On observing all the cases we can see that the proposed algorithm is most efficient in segmenting the tumor from the MRI. Tumor mapping, indicated by black outline, done on all the segmented images can be seen in Fig. 3.

To evaluate the effectiveness of IFCC algorithm we compare our results with other tumor detection algorithms from the literature. The proposed algorithm is compared with the state of the art methods including both automatic and semi-automatic segmentation methods as shown in Table 3. The comparison is done with prior information guided FCM(PIGFCM). LIPC based method, GWS, KIFCM, GLCM-CA, TLBO, AFBPNN and GRBF based segmentation. The results shows that the proposed IFCC algorithm performs better than other segmentation techniques and has lowest error percentage and highest DS and JS values.

The segmentation of brain tissues namely gray matter, white matter and cerebrospinal fluid is also carried out using the proposed technique by setting the value of parameter number of clusters equal to 3. The IFCC algorithm is run over several slices of both normal and autistic male and female dataset provided by NIMAC. In Table 4 and 5 patient numbers refers to following slices of dataset:P11–5074-004-02_10_t2_fit.nnrd (slice 100), P12–5074-004-02_10_t2_fit.nnrd (slice 140), P21–5128-004-01_10_T2_fit.nrrd (slice 80), P22–5128-004-01_10_T2_fit.nrrd (slice 114), P31–5150-004-02_10_T2_fit.nrrd (slice 110), P32–5150-004-02_10_T2_fit.nrrd (slice 67), P41–5157-004-02_10_T2_fit.nrrd (slice 70) and P42–5157-004-02_10_T2_fit.nrrd (slice 160).

The proposed approach for segmenting brain MRI images is also applied and validated for simulated 3D brain MRI images with varying level of noise from brain web database. The simulated dataset from brain web is provided by McGill University and can be obtained with different file extensions. (http://www.bic.mni.mcgill.ca/brainweb) It contains normal anatomical brain structures with size of each image is 181*217. The results are tested and validated for T1-weighted images with 3%, 5% noise levels and slice thickness of 1 mm (as shown in Table 6).

To evaluate the effectiveness of IFCC algorithm we compare our results with recent brain image segmentation algorithms from the literature as shown in Table 7. The proposed algorithm is compared with the state of the art methods namely, Local Information based IIFCM [21], RLS method [4], CNN [43] algorithm, NLSCHFCM [42] clustering algorithm, IFCM [17], RIFCM [28], BCSPNFCM [13], GRIFCM [13] and SFRCM [23] algorithm. Also, the comparison with the algorithms is done with varying levels of noise(3% and 5%).The results shows that the proposed IFCC algorithms performs better than other segmentation techniques and has highest value of DC. The average execution time in seconds of 50 runs of

Table 7 Comparative performance analysis of DC values for different brain tissue segmentation algorithms with varying noise level

			NOISE								
			0%			3%			5%		
AUTHOR	ALGORITHMS	Execution Time	GM	WM	CSF	GM	WM	CSF	GM	WM	CSF
Verma's [21]	IIFCM	14.51	0.881	0.842	0.846	0.854	0.792	0.724	0.832	0.757	0.711
Roy's [4]	RLS	15.99	0.891	0.864	0.850	0.834	0.803	0.794	0.812	0.773	0.729
Moeskops's [43]	CNN	55.10	0.923	0.891	0.882	0.919	0.871	0.811	0.894	0.800	0.782
Chen's [42]	NLSCHFCM	0.27	0.911	0.903	0.816	0.903	0.872	0.781	0.881	0.827	0.760
Kalaiselvi's [17]	IFCM	1.60	0.861	0.811	0.674	0.804	0.717	0.657	0.799	0.769	0.620
Dubey's [28]	RIFCM	21.40	0.911	0.900	0.825	0.890	0.863	0.756	0.872	0.829	0.727
Aparajeeta's [13]	BCSPNFCM	46.60	0.920	0.897	0.818	0.904	0.858	0.782	0.886	0.821	0.768
Namburu's [13]	GRIFCM	97.78	0.931	0.916	0.852	0.910	0.861	0.836	0.891	0.848	0.804
Namburu's [23]	SFRCM	85.20	0.947	0.918	0.851	0.892	0.848	0.829	0.886	0.818	0.810
Proposed	IFCC	18.50	0.985	0.964	0.900	0.970	0.948	0.846	0.957	0.943	0.850

these segmentation algorithms on real brain image is also given in Table 7. It can be observed that the execution time of IFCC is less than CNN,RIFCM, BCSPNFCM, GRIFCM, SFRCM and more than IIFCM, RLS, NLSCHFCM, IFCM algorithms.

5 Conclusion

An automatic brain tumor detection technique from real time MRI images has been presented in the paper. A novel intuitionistic and co-clustering based IFCC algorithm has been proposed. The Co-clustering approach offers the advantage of assigning membership functions for both object as well as features. The intuitionistic fuzzy set theory is found to be apt for handling uncertainty as compared to fuzzy sets theory. The proposed system is highly automatic as it produces tumor segmentation results in limited time and without any user intervention. The tumor detection result obtained by the proposed framework has been evaluated by experienced radiologist at Rajiv Gandhi Cancer Institute and Research Centre. Classification of three main tissues of brain is also performed. These tissues include- gray matter, white matter and cerebrospinal fluid. The performance of the proposed IFCC algorithms is evaluated in comparison with existing segmentation algorithms in the literature by performing experiments on real time database and simulated database with varying level of noise. The proposed algorithm is also compared on the basis of DC values with various brain image segmentation algorithms and is seen to outperform even in the presence of noise. The time taken by the proposed algorithm is greater than that of some of the earlier proposed work. However the performance of our algorithm is better than other segmentation algorithms. We aim to reduce the execution time in the future.

Acknowledgments The authors are thankful to the anonymous reviewers for their constructive suggestions to improve the overall quality of this paper. We would like to thank Dr. S. A. Rao, Radiologist, Rajiv Gandhi Cancer Institute & Research Centre, Delhi, India for guiding us about the identification of tumor and its proper evaluation. We would also like to thank RGCI & RC for providing the database so that we could carry out our research.

Compliance with ethical standards

Conflict of interest Author Heena Hooda declares that she has no conflict of interest. Author Om Prakash Verma declares that he has no conflict of interest.

References

1. Abdel-Maksoud E, Elmogy M, Al-Awadi R (2015) Brain tumor segmentation based on a hybrid clustering technique. Egyptian Informat J 16(1):71–81 ISSN 1110-8665
2. Amin J, Sharif M, Raza M, Yasmin M (2018) Detection of brain tumor based on features fusion and machine learning. J Ambient Intell Humaniz Comput. https://doi.org/10.1007/s12652-018-1092-9
3. Aparajeeta J, Nanda PK, Das N (2016) Modified possibilistic fuzzy C-means algorithms for segmentation of magnetic resonance image. Appl Soft Comput 41:104–119, ISSN 1568-4946. https://doi.org/10.1016/j.asoc.2015.12.003
4. Atanassov KT (1986) Intuitionistic fuzzy sets. Fuzzy Sets Syst 20(1):87–96
5. Benaichouche AN, Oulhadj H, Siarry P (2013) Improved spatial fuzzy c-means clustering for image segmentation using PSO initialization, Mahalanobis distance and post-segmentation correction. Digital Signal Process 23(5):1390–1400, ISSN 1051-2004. https://doi.org/10.1016/j.dsp.2013.07.005

6. S. M. Bhandarkar, Nammalwar, 'segmentation of multispectral MR images using a hierarchical self-organizing map', proceedings of the 14th IEEE symposium on computer-based medical system CBMS, Vol: 26, No: 27, pp: 294–299,2001.
7. Bricq S, Collet C, Armspach JP (2008) Unifying framework for multimodal brain MRI segmentation based on Hidden Markov Chains. Medical Image Analysis 12(6):639–652
8. Caldairou B, Passat N, Habas PA (2011) A non-local fuzzy segmentation method: application to brain MRI. Biomed Image Comput Group, Pattern Recognition 44:1916–1927
9. Charles R. Noback, Norman L. Strominger, Robert J. Demarest and David A. Ruggiero: The Human Nervous System: Structure and Function, 6th ed., Humana Press, 2005.
10. Chen Y, Zhang J, Wang S, Zheng Y (2012) Brain magnetic resonance image segmentation based on an adapted non-local fuzzy c-means method. IET Computer Vision 6(6):610–625. https://doi.org/10.1049/iet-cvi.2011.0263
11. Chen Y, Li J, Zhang H, Zheng Y, Jeon B, Wu QJ (2016) Non-local-based spatially constrained hierarchical fuzzy C-means method for brain magnetic resonance imaging segmentation. *IET Image Processing* 10(11): 865–876. https://doi.org/10.1049/iet-ipr.2016.0271
12. Deepa AR, Sam Emmanuel W (2018) An efficient detection of brain tumor using fused feature adaptive firefly backpropagation neural network. Multimed Tools Appl 78:11799–11814. https://doi.org/10.1007/s11042-018-6731-9
13. Dice LR (1945) Measures of the amount of ecologic association between species. Ecology 26:297–302
14. Dubey YK, Mushrif MM, Mitra K (2016) Segmentation of brain MR images using rough set based intuitionistic fuzzy clustering. Biocybernet Biomed Eng 36(2):413–426, ISSN 0208-5216. https://doi.org/10.1016/j.bbe.2016.01.001
15. El-Melegy, Moumen T, Mokhtar, Hashim M., 'Tumor segmentation in brain MRI using a fuzzy approach with class center priors', EURASIP Journal on Image and Video Processing, 2014, 2014:21.
16. Yong Fan, Hengyi Rao, Joan Giannetta, Hallam Hurt, Jiongjiong Wang, Christos Davatzikos, and Dinggang Shen, 'Diagnosis of brain abnormality using both structural and functional MR images', in Proceedings of the IEEE conference on Eng med BioI Soc., 2006
17. Madasu Hanmandlu, Om Prakash Verma, Seba Susan, V.K. Madasu, 'Color Segmentation by Fuzzy co-clustering color features', Neurocomput, volume 120, 2013.
18. Hanuman Verma RK, Agrawal AS (2016) An improved intuitionistic fuzzy c-means clustering algorithm incorporating local information for brain image segmentation. Appl Soft Comput 46:543–557, ISSN 1568-4946. https://doi.org/10.1016/j.asoc.2015.12.022
19. H. Hooda, O. P. Verma and T. Singhal, 'Brain tumor segmentation: A performance analysis using K-means, fuzzy C-means and region growing algorithm', IEEE International Conference on Advanced Communications, Control and Computing Technologies, Ramanathapuram, 2014, pp. 1621–1626.
20. Huang M, Yang W, Wu Y, Jiang J, Chen W, Feng Q (2014) Brain tumor segmentation based on local independent projection-based classification. IEEE Trans Biomed Eng 61(10):2633–2645. https://doi.org/10.1109/TBME.2014.2325410
21. Hyung LK, Song YS, Lee KM (1994) Similarity measure between fuzzy sets andbetween elements. Fuzzy Sets Syst 62(3):291–293
22. K. M. Iftekharuddin, 'On techniques in fractal analysis and their applications in brain', Medical imaging systems: technology and applications, Analysis and Computational Methods, vol. 1, World Scientific Publications, ISBN 981–256–993-6, 2005.
23. Jaccard P (1912) The distribution of flora in the alpine zone. New Phytol 11:37–50
24. Zexuan Ji, Yong Xia: Fuzzy Local Gaussian Mixture Model for brain MR Image Segmentation, IEEE transactions on Information Technology in Biomedicine, VOL.16 NO.3, 2012
25. T. Kalaiselvi P, P. Nagaraja P, V. Ganapathy Karthick, "Improved Fuzzy C-Means for Brain Tissue Segmentation Using T1- Weighted MRI Head Scans", Int J Innovativ Sci, Eng Technol, Vol. 3 Issue 7, 2016.
26. Kaus MR, Warfield SK, Nabavi A, Black PM, Jolesz FA, Kikinis R (2001) Automated segmentation of MR images of brain tumors. Radiology 218:586–591
27. A.R. Kavitha, C. Chellamuthu, Kavin Rupa, 'An Efficient Approach for Brain Tumor Detection Based on Modified Region Growing and Neural Network in MRI Images', IEEE International Conference on Computing, Electronics and Electrical Technologies, pp. 1087–1095, 2012.
28. Kennedy J, Eberhart RC (1995) Particle swarm optimization. Proc IEEE Int Conf Neural Netw:1942–1948
29. Mahmood Q, Chodorowski A, Persson M (2015) Automated MRI brain tissue segmentation based on mean shift and fuzzy c-means using a priori tissue probability maps. IRBM 36(3):185–196, ISSN 1959-0318. https://doi.org/10.1016/j.irbm.2015.01.007
30. Wahba Marian: An automated modified region growing technique for prostate segmentation in trans-rectal ultrasound images, Master's Thesis, Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, Canada, 2008.

31. Mayer A, Greenspan H (2009) An adaptive mean-shift framework for MRI brain segmentation. IEEE Trans Med Imaging 28(8):1238–1250
32. Moeskops P, Viergever MA, Mendrik AM, de Vries LS, Benders MJNL, Išgum I (2016) Automatic segmentation of MR brain images with a convolutional neural network. IEEE Trans Med Imaging 35(5): 1252–1261. https://doi.org/10.1109/TMI.2016.2548501
33. Nabizadeh N, Kubat M (2015) Brain tumors detection and segmentation in MR images: Gabor wavelet vs. statistical features. Comput Electr Eng 45:286–301. https://doi.org/10.1016/j.compeleceng.2015.02.007
34. Nakib A, Oulhadj H, Siarry P (2009) A thresholding method based on two-dimensional fractional differentiation. Image Vis Comput 27:1343–1357
35. Namburu A, Samay SK (2017) Srinivasa Reddy Edara, Soft fuzzy rough set-based MR brain image segmentation. Appl Soft Comput 54:456–466, ISSN 1568-4946. https://doi.org/10.1016/j.asoc.2016.08.020
36. Namburu A, Samayamantula SK, Edara SR Generalised rough intuitionistic fuzzy c-means for magnetic resonance brain image segmentation. IET Image Process 11(9):777–785
37. Ortiz A, Górriz JM, Ramírez J, Salas-Gonzalez D (2011) MR brain image segmentation by growing hierarchical SOM and probability clustering. Electronics Lett 47(10):585–586. https://doi.org/10.1049/el.2011.0322
38. Prastawa M (2004) E Bullitt, S ho, G Gerig, 'A brain tumor segmentation framework based on outlier detection'. Med Image Anal 8:275–283
39. Rajinikanth V, Satapathy SC, Fernandes SL, Nachiappan S (2017) Entropy based segmentation of tumor from brain MR images – a study with teaching learning based optimization. Pattern Recognition Lett 94:87–95, ISSN 0167-8655. https://doi.org/10.1016/j.patrec.2017.05.028
40. Roy S, Bandyopadhyay SK (2016) A New Method of Brain Tissues Segmentation from MRI with Accuracy Estimation. Procedia Comput Sci 85:362–369, ISSN 1877-0509. https://doi.org/10.1016/j.procs.2016.05.244
41. Sompong C, Wongthanavasu S (2016) An efficient brain tumor segmentation based on cellular automata and improved tumor-cut algorithm. Expert Syst Appl 72:231–244. https://doi.org/10.1016/j.eswa.2016.10.064
42. M. Sugeno, Fuzzy Measures and Fuzzy Integrals – A Survey, North-Holland,1977.[26] R.R. Yager, On the measure of fuzziness and negation. Part I: Membership inthe unit interval, Int J Gen Syst (1979) 221–229.
43. Xu Z (2007) Some similarity measures of intuitionistic fuzzy sets and their appli-cations to multiple attribute decision making. Fuzzy Optim Decis Mak 6(2):109–121
44. Zadeh LA (1965) Fuzzy sets. Inf Control 8(3):338–353

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Om Prakash Verma is Principal of GB Pant Govt. Engineering College. Prior to this, he was head at Department of CSE, Delhi Technological University,Delhi. He has a wide professional experience of 25 years along with distinguished research activities. His current research area includes softcomputing,evolutionary algorithms and many more.In addition to a successful academician he is a distinguished researcher. He has authored more than 70 publication in International Journal and Conferences of repute. He has supervised more than 35 M-tech dissertations. Currently, 8 Ph.D thesis are in progress under his supervision and 3 has already completed.



Heena Hooda recieved the Masters degree from Delhi Technological University, Delhi, India in 2013. She is currently working towards the Ph. D degree from Department of Information Technology, Delhi Technological University, Delhi, India. Her area of interests include, artificial intelligence, soft computing, pattern recognition and medical image processing.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/342389062

A Novel MS-Word Text Document Multiple Watermarking Technique for Authentication and Copyright Protection

Article *in* International Journal of Engineering Trends and Technology · May 2020

DOI: 10.14445/22315381/IJETT-V68I5P201S

CITATIONS
0

READS
32

3 authors, including:



Jeebananda Panda
Delhi Technological University
14 PUBLICATIONS **28** CITATIONS

SEE PROFILE

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

watermarking of audio View project

All content following this page was uploaded by Jeebananda Panda on 18 August 2020.

The user has requested enhancement of the downloaded file.

A Novel MS-Word Text Document Multiple Watermarking Technique for Authentication and Copyright Protection

Ayush Arora[#1], Drishti Agarwal[#2], Jeebananda Panda[#3]

[#1, 2, 3]*Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India*

Abstract*—In today's time when everyone is surrounded by digital data, the development of internet-related technologies, like cloud databases, social media platforms and much more, is rapidly increasing which raises the problem of information security. Today anyone can easily generate identical copies of any text document. So, to protect the data or the content i.e., intellectual property of someone, some measures are needed to verify the authenticity of one's work. Copyright protection is one of the most important and difficult challenges for the researchers. So, for verification, many researchers have proposed several algorithms to embed a watermark in a text document (In simple words embedding a watermark means hiding some secret information in the document which is hard to detect). The proposed approach uses the application class property of MS-Word document, RGB colour values of text and spacing between the lines to hide the watermark in a monocoloured MS-Word text document. Several different documents are used for evaluation which got attacked by different types of changes.*

Keywords — *MS-Word Document, watermarking, copyright, vowels, security.*

I. INTRODUCTION

Nowadays, the digitalisation of information is becoming more and more common. In this modern digital world, the spread of the information in the digital format is rapidly increasing. Advancement in internet technologies is one of the factors that encourages this increase. Various forms of data such as audio, video, text and images can be quickly communicated between two points in this era. With this ease comes the concern of securing this information from unauthorized users and protecting the authenticity of the digital work. There are a number of ways which are used by hackers today to retrieve the data making it susceptible to manipulation and fraud. One of the solutions to this problem is digital watermarking which hides a secret message into the digital content and hence plays a vital role in copyright protection and authentication of the content. The idea is to place a secret message inside the text document without losing the meaning of the digital content which later on can be used for verification of ownership. The proposed method uses multiple levels of text watermarking for the same.

In the current pandemic scenario of COVID 19, mostly all schools and universities will have to take online examinations to cope up with the same. Hence, there would be an increase in the digital content. Watermarking the answer sheets of students would help them claim their authenticity. Since the proposed algorithm gives an easy and fast watermark embedding way, hence the algorithm is very useful in the current scenario.

Every organisation, small or big, has certain important text documents associated with it which are integral to its working. The problem with the existing techniques is that they are not robust enough and less secure. The challenge is two-folded. First is to detect the presence of any formatting attacks, and second is to prove ownership of the document. A new framework is proposed here which deals with these challenges and is more robust, imperceptible and easy to embed.

Main points covering contribution to this research are:

1. The work proposes a new digital text watermarking technique which uses certain properties of both application class and document class of a monocoloured MS Word Document to embed a watermark.
2. Apart from these properties, the count of certain alphabets is used for embedding the watermark thus ensuring a quick detection of any kind of formatting attack.
3. The proposed work is robustup to 99.99% against formatting attacks and is highly secured.
4. Any attacker would be unable to remove the watermark without strongly altering the watermarked document.

The rest of the paper is structured as follows: Section 2 shows the recent and related work done in the field. Section 3 deals with the detailed description of the proposed method followed by section 4 where results obtained are discussed. Section 5 concludes the paper and gives direction to the future work.

II. RELATED WORK

Watermarking in text documents is an active area of research. There is a lot of work that has been done in watermarking of audio, video and image files[1] [2]. In case of text documents some interesting work has been done by many scholars using various approaches. Kim *et al.* [3] has proposed a technique in which they grouped adjacent words to form segments and then classified the segments using word class information, then they hide the information in these segments using spacing between the words of the same class whereas Alotaibi *et al.* [4] has used the pseudo-space that is very small space between two parts of the same word they have altered this space to store the binary bits in the document. Jaman*et al.* [5] proposed an algorithm using natural language processing, and in their work, they have used various grammatical rules like verbs, conjunction, preposition, modals and articles for watermarking [6]. Some scholars have used space between words and lines to hide watermark [7] [8][9].

Iqbal *et al.* [10] uses the special properties of MS-Word documents that cannot be altered by any formatting or change in content of the document [11] [12], in their approach they hide the encrypted key to insert in these special properties. Khadam *et al.* [13] has also used these special properties to hide they watermark, in their proposed work they have generated an encrypted message, then they convert them to a binary string followed by converting to a decimal string, after which they divide that string to four equal parts followed by taking logarithms of these number, whose value they insert in the top, right, left and bottom margin of the MS-Word document.

Rizzo *et al.* [14] has proposed a technique in which they substitute Latin symbols with homoglyph symbols to prevent any visual change in the content of the text document they have used different Unicode symbols for substitution which looks nearly the same as of the symbol to be substituted, in proposed work they embed their secret password by substitution of the Latin symbols by homoglyph symbols. [15] glyphs are also used to hide the watermark.

Afify*et al.* [16] has proposed a hybrid approach which they have used the concept of natural language processing along with structural technique and hashing. They have used the most occurring word in the document to embed their watermark in the document. They have altered the space in the left of the word to insert a bit in that word along with this watermark. They have also used MD5 hashing technique to hash the complete document to check if the document has been tampered or not.

III. THE PROPOSED METHOD

In this section, the detailed algorithm and framework used for embedding watermarks in a text document is presented followed by its extraction. The document class of MS Word has many properties (WSP) which can be used to store the watermark without actually affecting the contents of the document. Some of these properties have been exploited to embed the multi-level watermark. The proposed technique is highly robust. The entire procedure is divided into three sections: The first section deals with the method of generating watermark followed by the embodiment of this watermark in the document and then finally checking against any formatting attack by extracting the watermark. Figure 1 shows the complete architecture of the text watermark system.

A. Generating Watermark:

This section describes how the watermarks to be embedded in the document are generated. Three different types of watermarks are generated. For the first type, the aim was to embed a part of the generated watermark in as many words as possible. For this, research was done and it was found that in most cases 100% of the document can be covered if the watermark is hidden using vowels and the letter ‘y’. Algorithm 1 is used for the same. The entire document is iterated over and the individual counts of all the five vowels along with the letter ‘y’ is stored. Binary strings corresponding to these counts are randomly generated. These binary strings serve the purpose of the hidden information that has to be embedded in the text document. The second type of watermark is generated by taking in account the number of lines in the document. A binary string of length equal to the number of lines in the document is generated and then embedded. The third watermark will be an encrypted key. Figure 2 shows the generation of watermark for a text document and Algorithm 1 describes the pseudo code for the same.

B. Embedding Watermark:

In the proposed approach, multi-level embedding of watermarks is being performed. The watermarks generated in part A are to be embedded in the text document. This is done using two properties of the document class. Algorithm 2 describes the entire procedure for embedding watermark.

The watermarks generated against the count of the vowels and the letter ‘y’ are stored in the document by using the font colour of the document. Each of the values 0 and 1 are assigned a particular font colour value. As the document is iterated over, the font colour of the vowels and the letter ‘y’ are changed as per the corresponding character of the binary string. This process is followed for all the six letters under consideration. Thus, using the font

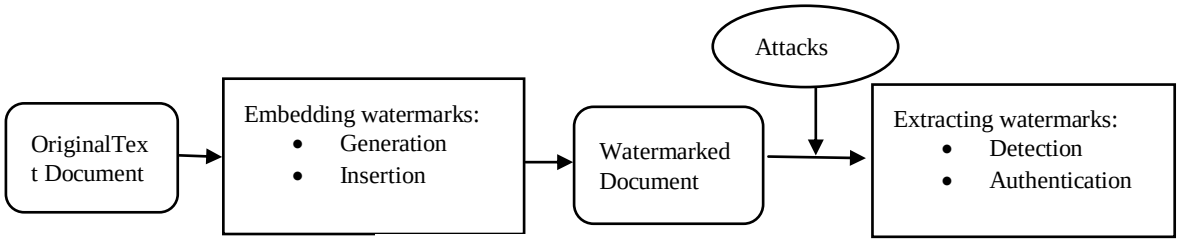


Figure 1: Architecture of TextWatermark

colour of the text document six watermarks are embedded.One of the major advantages of applying this technique is that it covers almost every word of the document, hence even a minor change in the content of the document would be easily detectable.

Algorithm 1: Watermark Generation

```
Input: MS-Word Doc (T)
Output: s_a, s_e, s_i, s_o, s_u, s_y, s_line // Secret binary string
Start:
Data: c_a, c_e, c_i, c_o, c_u, c_y // Count of character a, e, i, o, u, y
Variable Declaration:
        num_char = Number of characters in T
        c_line = number of lines in T
Initialization:
        c_a = 0, c_e = 0, c_i = 0, c_o = 0, c_u = 0, c_y = 0
        for(i = 1 to num_char) do
            curr_char = T[i]
            if(curr_char == 'a' or curr_char == 'A') then
                c_a = c_a + 1
            elif(curr_char == 'e' or curr_char == 'E') then
                c_e = c_e + 1
            elif(curr_char == 'i' or curr_char == 'I') then
                c_i = c_i + 1
            elif (curr_char == 'o' or curr_char == 'O') then
                c_o = c_o + 1
            elif(curr_char == 'u' or curr_char == 'U') then
                c_u = c_u + 1
            elif(curr_char == 'y' or curr_char == 'Y') then
                c_y = c_y + 1
            end if
        end for
        s_a <- Random binary string of length c_a
        s_e <- Random binary string of length c_e
        s_i <- Random binary string of length c_i
        s_o <- Random binary string of length c_o
        s_u <- Random binary string of length c_u
        s_y <- Random binary string of length c_y
        s_line <- Random binary string of length c_line

    end
```

To ensure extra security, the second type of generated watermark is embedded in the document using another property of the document class which is the space between the lines. Again 0 and 1 are assigned different values and the entire document is iterated over each line. Corresponding to the value of the character in the string, space between the lines is altered.This second type of watermark embedding still allows to prove ownership of the digital data even if some formatting attack changes the font colour of the document. The third kind of watermark which is inserted is the encrypted key. It is inserted using the application class of MS-Word document and remains unaffected by any change in the content and formatting of the document.Apart from ownership verification, this step becomes crucially important in the case when some formatting attack alters both the font colour and spacing between lines of the entire document. After the watermarks are embedded, the watermarked document is in the MS Word format only.

Algorithm 2: Watermark Embedding

```
Input: MS-Word Doc (T), s_a, s_e, s_i, s_o, s_u, s_y, s_line, Secret Key (SK)
Output: Watermarked Document(T')
Start:
Data:
        RGB_1, RGB_2 // RGB color values
        Space_1, Space_2 // Line space values
Variable Declaration:
        num_char = Number of characters in T
        c_line = number of lines in T
        WSP = MS-Word Special Property
Initialization:
        i_a = 0, i_e = 0, i_i = 0, i_o = 0, i_u = 0, i_y = 0
        WSP = SK
        for(i = 1 to num_char) do
            curr_char = T[i]
            if(curr_char == x) then // where x is a, e, i, o, u, y
                if(s_x[i_x] == '0') then
                    T[i].Color = RGB_1
                else
                    T[i].Color = RGB_2
                end if
                i_x = i_x + 1
            end if
        end for
        for(i = 1 to c_line) do
            if(s_line == '0') then
                T.Line[i].Space = Space_1
            else
                T.Line[i].Space = Space_2
            end if
        end for

    end
```

C. Watermark Extraction:

Watermark extraction is the reverse process of embedding watermark. It is the step where any possible format attacks are detected and ownership of the document is proved. The hidden watermark is extracted from the text document and compared with

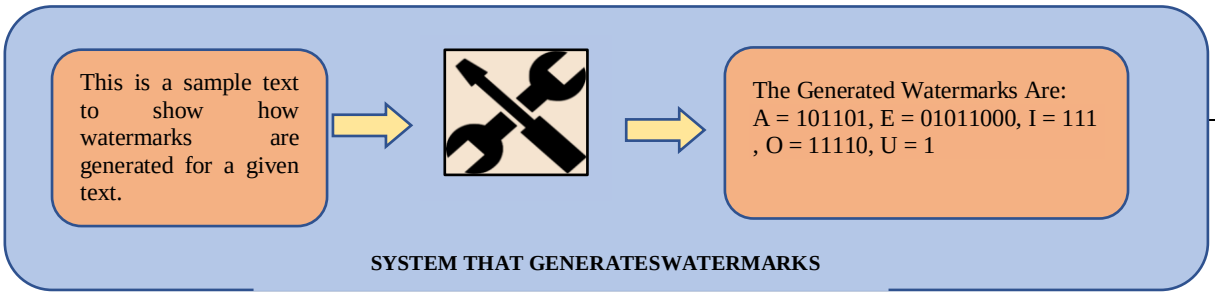


Figure 2: Generating Watermarks

the embedded watermark for the same. The detailed process is described in Algorithm 3.

The system iterates over the entiredocument and stores the colour code of the five vowels and the letter 'y' in the form of binary strings. Hence six binary strings are generated from the document which are then compared with the original binary strings that were embedded in the document.Any kind of change in the content of the document generates a different binary key of one or more letters, thus marking the presence of some formatting attack. Second type of watermark is extracted by iterating over all the lines of the document. Again, a binary string is generated from the input document depending on the value of the line spaces which is then compared with the string embedded in the document. Third type of watermark is extracted from the application class property of MS-Word document. This helps the user acclaim his/her ownership of the document.

Algorithm 3: Watermark Extraction

```
Input: MS-Word Doc (T')
Output: s_a, s_e, s_i, s_o, s_u, s_y, s_line, SK
Start:
Data: RGB_1, RGB_2 // RGB color values
        Space_1, Space_2 // Line space values
Variable Declaration:
        num_char = Number of characters in T'
        c_line = number of lines in T'
        WSP = MS-Word Special Property
Initialization:
        i_a = 0, i_e = 0, i_i = 0, i_o = 0, i_u = 0, i_y = 0
        SK = WSP
          for(i = 1 to num_char) do
            curr_char = T'[i]
            if(curr_char == x) then // where x is a, e, i, o, u, y
              if(T'[i].Color == RGB_1) then
                s_x += '0'
              elif(T'[i].Color = RGB_2) then
                s_x += '1'
        else
                throw error('Document is tempered')
              end if
            end if
          end for
          for(i = 1 to c_line) do
            if(T.Line[i].Space = Space_1) then
              s_line += '0'
            elif(T.Line[i].Space = Space_2) then
              s_line += '1'
            else
```

IV. RESULTS AND DISCUSSION

Multiple experiments were performed to test the system. Following two parameters were used as the basis of evaluation:

1. Robustness
2. Imperceptibility

The first two types of watermark embedded are text dependent whereas the third type of watermark that is inserted for the purpose of experiments is "This is a secret message."

A. Robustness

Any document is subject to several types of formatting attacks. A system is said to be robust if it is able to recover its embedded watermark even after such attacks. The multilevel watermarking done helps achieve a very high rate of robustness because even if the first two watermarks are lost, the third one is recovered as it is. The two parameters used to compute robustness are Pattern matching rate (PMR) and Watermark Distortion Rate (WDR) and are stated as follows:

$$PMR = \frac{No.\,of\ patterns\ matched\ correctly}{No.\,of\ watermark\ patterns}$$

$$WDR = 1 - PMR$$

The proposed technique is 99.99% robust against formatting attacks. Table I shows the robustness value for various documents. The table values were obtained after applying all types of formatting attacks to the documents. Formatting attacks like bold, italics, underline, font size, font type were applied and it was found that all types of watermarks were recovered with 100% accuracy. A message showing "First two watermarks recovered. Third watermark is: This is a secret message" pops up. However, change in font colour destroys one of the watermarks and change in spacing between lines removes the other type of watermark. Yet the third type of watermark was always recovered making the system highly robust against formatting attacks. Figure 3 shows two cases: one is when all three watermarks are recovered and the other is when one of the first two watermarks are destroyed but third is recovered as it is.

.

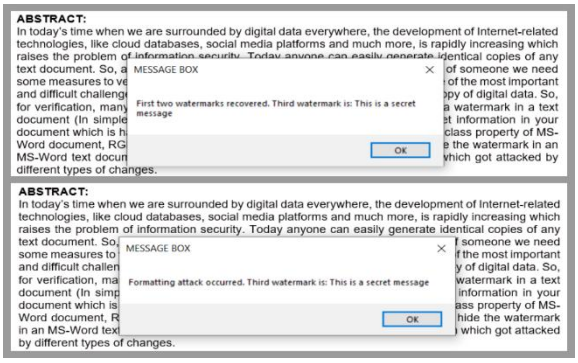


Figure 3: Top Figure: When all three Watermarks are Recovered. Bottom figure: When one of the First Two Watermarks is Lost but Third is Recovered as it is.

Table I: Robustness Achieved for Various Documents

DOCUMENT	ROBUSTNESS ACHIEVED
Document 1	100%
Document 2	100%
Document 3	99.99%

B. Imperceptibility

A system is said to be imperceptible if it allows embedding watermarks in a text document without affecting the contents of the documents. It is both a primary and fundamental requirement of the watermark. The two measures used to measure imperceptibility are Peak Signal to Noise Ratio (PSNR) and Similarity (SIM) percentage which are formalized as follows:

$$PSNR = 20log_{10}\frac{O_{doc}(Max)}{RMSE}$$

$$SIM = [1 - \frac{RMSE}{O_{doc}(Max)}] \times 100$$

where O_{doc}(Max) is a maximum pixel value in a document and RMSE (root mean squared error) is calculated as follows:

$$RMSE = \sqrt{\frac{1}{m \times n}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[O_{doc}(i.j) - W_{doc}(i,j)]^2}$$

Figure 4 shows the images of the document before and after embedding the watermarks. The images can be seen to be similar showing the high imperceptibility of the system. Extensive experiments were done to measure PSNR and SIM values. To calculate the PSNR and SIM values a python system has been used in which the original image and the watermarked image of the document are loaded as a matrix by OpenCV. O_{doc}(max) value is the maximum value of the matrix. NumPy is used to find the RMSE value followed by PSNR and SIM values.The PSNR values were found to be between 33.499 and 45.821. And SIM values were found between 97.886 and 99.2. The acceptable value of PSNR should be above 30 to attain a high level of imperceptibility. Table II shows PSNR and SIM values for different documents.

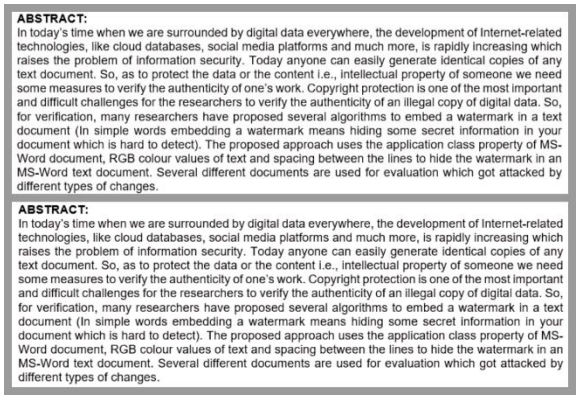


Figure 4: Top figure: Original document. Bottom figure: Watermarked Document.

Table II: PSNR and SIM Values for Different Documents

DOCUMENT	PSNR	SIM
Document 1	33.499	97.886
Document 2	38.467	98.13
Document 3	45.821	99.2

V. CONCLUSION AND FUTURE WORK

A robust and secure method is proposed to authenticate the digital contents as the system embeds multiple watermarks. Several techniques are proposed in this field. But the proposed system achieves higher levels of security as it uses multiple level of watermarks. Through experiments it was seen that the proposed work is low in capacity but highly robust, imperceptible and also fast in embedding of watermark. In the future, the work will be extended for detection of all types of the formatting attacks and location of changes in the document. Apart from detection, it is proposed to come up with a system that can get back the original document. And specifically, for online examination due to COVID 19 pandemic scenario, watermarking can be performed on images of diagrams provided by the studentto claim their authenticity.

REFERENCES

[1] Mayer, Joceli, Paulo VK Borges, and Steven J. Simske. "*Text Watermarking.*" In Fundamentals and Applications of Hardcopy Communication, pp. 43-113. Springer, Cham, 2018.

[2] Nematollahi, Mohammad Ali, Chalee Vorakulpipat, and Hamurabi Gamboa Rosales. "*Text Watermarking.*" In *Digital Watermarking*, pp. 121-129. Springer, Singapore, 2017.

[3] Kim, Young-Won, Kyung-Ae Moon, and Il-Seok Oh. "*A Text Watermarking Algorithm based on Word Classification and Inter-word Space Statistics*." In *ICDAR*, pp. 775-779. 2003.

[4] Alotaibi, Reem A., and Lamiaa A. Elrefaei. "*Utilizing word space with pointed and un-pointed letters for Arabic text watermarking.*" In 2016 UKSim-AMSS 18th International Conference on Computer Modelling and Simulation (UKSim), pp. 111-116. IEEE, 2016.

[5] Jaman, Khandokar Nafis, Zannatun Nayem, Bristi Rani Roy, Nayreet Islam, Faisal R. Badal, and Subrata K. Sarker. "*A Text Watermarking Algorithm Developed Using Natural Language Processing.*" (2019).

[6] Yingjie, Meng, Liu Huiran, Shang Tong, and Teng Xiaoyu. "*A zero-watermarking scheme for prose writings*." In 2017 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), pp. 276-282. IEEE, 2017.

[7] Kuribayashi, Minoru, Takuya Fukushima, and Nobuo Funabiki. "*Robust and Secure Data Hiding for PDF Text Document.*" IEICE TRANSACTIONS on Information and Systems 102, no. 1 (2019): 41-47.

[8] Taha, Ahmed, Aya S. Hammad, and Mazen M. Selim. "*A high capacity algorithm for information hiding in Arabic text.*" Journal of King Saud University-Computer and Information Sciences (2018).

[9] Liang, Ooi Wei, and VahabIranmanesh. *"Information hiding using whitespace technique in Microsoft word."* In 2016 22nd International Conference on Virtual System & Multimedia (VSMM), pp. 1-5. IEEE, 2016.

[10] Iqbal, Muhammad Munwar, Umair Khadam, Ki Jun Han, Jihun Han, and Sohail Jabbar. "*A Robust Digital Watermarking Algorithm for Text Document Copyright Protection based on Feature Coding*." In *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, pp. 1940-1945. IEEE, 2019.

[11] Ghafir, Ibrahim, Jibran Saleem, Mohammad Hammoudeh, Hanan Faour, Vaclav Prenosil, Sardar Jaf, Sohail Jabbar, and Thar Baker. "*Security threats to critical infrastructure: the human factor*." The Journal of Supercomputing 74, no. 10 (2018): 4986-5002.

[12] Muzammal, Syeda Mariam, Munam Ali Shah, Hasan Ali Khattak, Sohail Jabbar, Ghufran Ahmed, Shehzad Khalid, Shahid Hussain, and Kijun Han. *"Counter measuring conceivable security threats on smart healthcare devices."* IEEE Access 6 (2018): 20722-20733.

[13] Khadam, Umair, Muhammad Munwar Iqbal, Muhammad Awais Azam, Shehzad Khalid, Seungmin Rho, and Naveen Chilamkurti. "*Digital Watermarking Technique for Text Document Protection Using Data Mining Analysis*." *IEEE Access* 7 (2019): 64955-64965.

[14] Rizzo, Stefano Giovanni, Flavio Bertini, and Danilo Montesi. "*Content-preserving text watermarking through unicode homoglyph substitution*." In Proceedings of the 20th International Database Engineering & Applications Symposium, pp. 97-104. 2016.

[15] Xiao, Chang, Cheng Zhang, and Changxi Zheng. "*Font Code: Embedding information in text documents using glyph perturbation.*" ACM Transactions on Graphics (TOG) 37, no. 2 (2018): 1-16.

[16] Afify, Amira Eid, Ahmed Emran, and Ahmed Yahya. *"A Tamper proofing text watermarking shift algorithm for copyright protection*." Arab Journal of Nuclear Sciences and Applications 52, no. 3 (2019): 126-133.

View publication stats

Proceedings of the Second International Conference on Inventive Research in Computing Applications (ICIRCA-2020)
IEEE Xplore Part Number: CFP20N67-ART; ISBN: 978-1-7281-5374-2

A Route Stable Energy and Mobility aware routing protocol for IoT

Rajesh Kumar Yadav
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
rkyadav@dtu.ac.in

Nimisha Awasthi
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
awasthi.nimisha97@gmial.com

***Abstract*-Mobility enhancement in Low power device has always been a great deal. With the advancement in The Internet of Things (IoT), mobility has become one of its characteristic. Routing in IoT is challenging and mobility even complicates it more. The Routing over Low Power and Lossy Networks (ROLL) group at Internet Engineering Task Force (IEFT) developed a routing protocol for some IoT devices called RPL, a protocol for communication among low power devices in a lossy network. Through RPL protocol has enhanced the routing in Low power lossy networks but it does not address the issue of mobility, it is mostly for static networks. A decent number of protocols are developed for integrating mobility in RPL. One of them is EMA-RPL. The issue with EMA-RPL is that the routes are unstable and the preferred parent of mobile node sometimes changes frequently. This work enhances the EMA-RPL protocol by selecting the preferred parent of a mobile node which last for a considerable amount of time and hence improving route stability. The proposed method observe the speed of the mobile node and accordingly decide which node can be selected as a parent node for the mobile node.**

Keywords— RPL, Mobility, Handover delay, Route Stability.

I INTRODUCTION

The internet is growing exponentially and is covering all the day-to-day life areas possible. The upgraded version of the Internet is IoT(Internet of Things)[1,2] which is now bringing a revolution in technology. The pace at which the Internet of Things is growing has given many opportunities to many organizations to provide specification and guideline for IoT devices. A large number of things are now part of IoT. Some of the major areas are RFID(tags), wireless sensor networks, Near field communication (NFC). Wireless sensor network (WSN)[3] is a group of resource-constrained devices called sensors for monitoring and recording the physical conditions of the environmental surroundings and send the collected data at a central location. WSN cooperatively pass data to a border router or sink where the data is sent to other networks for analyzing. A sink or border router acts as an interface between users and the network. A wireless sensor network contains a large number of sensor nodes equipped with sensing and computing devices, radio transceivers and power components. Wireless sensor networks have gained immense popularity due to their flexibility in solving problems in different application domains and have the potential to change our lives. WSNs have been successfully applied in various application domains like military, area monitoring, transportation, health applications[4] etc.

The problem is the nodes in wireless sensor network (WSN) are resource constrained, they have limited processing speed, storage capacity, power supply and communication bandwidth[5]. These type of resource constrained networks are called Low power and Lossy Networks (LLN)[6] and routing in these types of the network brings in new challenges. In Low power and Lossy Networks (LLNs) both the routers and their interconnect are constrained. LLN routers typically operate with constraints and their interconnects are characterized by high loss and, low data rates and instability. The IETF developed the IPv6 low-power wireless personal area networks (6LoWPAN)[7,8,9] which is used as an adaptation layer that allows sensor nodes to implement the Internet protocol (IP) stack and become accessible by other devices on the network. This new layer allows sensor nodes to implement routing protocols at the network layer and provide connectivity that enables countless applications. With the diverse immense number of IoT devices, such a change in technology cannot be handled by conventional routing protocols. They no longer can accommodate a large number of added nodes. For this reason, RPL[10] was designed especially for LLNs to tackle the problem of routing.

Mobility[11] in IoT brings a new revolutionary change in the technology, enhancing the flexibility of a network, enhanced and innovative products and enlarging IoT to as many fields as possible. With advancement comes in the new challenges, mobility brings an issue that needs to be dealt with. RPL is basically for static networks which do not handle mobility. As such, it is unable to adequately deal with some mobility problems. An inconsistency results from a node failure or a physical movement causing a change in the topology. In the RPL[12] Standard specification, due to mobility, the node becomes detached from all its connected nodes and loses its data. Consequently, it finds a new routing path and updates its value. RPL deals with mobility with a technique called self-healing strategy [13], but it does not provide the desired optimal result. As solution protocols were developed to address the issue of mobility. A large number of researchers

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:21:15 UTC from IEEE Xplore. Restrictions apply.

have worked upon the issue of mobility. This paper introduces those works and compares and analyze various mobility enhanced protocol on various parameters.

Implementing mobility in RPL arises new issues. The first thing that needs to be looked after is the data loss. Once a connection is lost and if this loss in connection is not known at an earlier stage then data loss occurs. To avoid this data loss many protocols use a proactive approach which detects the mobility before any node gets disconnected from its parent node. In these types of approaches if the system identifies that a node might get disconnected it notify about this to the mobile node and the mobile node stops sending the data until a new parent node is identified. This is called the handover delay. The minimum the handover delay the better the results.

Another deal-breaking thing is route stability. If an unstable route is selected then it causes a lot of flow of control messages and frequent change in the topology of the network which gives undesirable results. In this paper, the route stability is considered as the main goal and an algorithm is presented to improve it.

The paper is organized in the following structure: Section (II) gives some of the work performed by other authors. Section (III) presents the proposed protocol A route stable energy and mobility aware routing protocol. Performance Evaluation is present in Section (IV) and section (V) concludes the paper.

II RELATED WORK

This section specifies the work of some other authors who have integrated mobility in RPL.

In Mobility Enhanced RPL(ME-RPL) [14], the author resolves the mobility issue by not allowing a mobile node to be a parent node. All the mobile nodes are leaf node in DODAG, hence route stability is achieved. To avoid mobile node disconnection it uses a DIS timer that changes according to the frequency of change of preferred parent. Hence triggering DIO message before trickle timer. Though ME-RPL resolves mobility in RPL to a great extent it does not consider the high energy consumption due increase in control message when DIS timer is decreased. It also does not consider the handover delay[15]. This protocol also gives no information on how the data will be handled during mobility.

In [16] Energy and Mobility Aware RPL (EMA-RPL), the author considers the same conditions as that of ME-RPL that the mobile node will be a leaf node, but unlike ME-RPL it is a proactive protocol which before the disconnection of mobile node make it aware of its mobility using RSSI(Received signal strength indicator)[17] value. As soon as RSSI value is reduced below a threshold the parent node informs this to the mobile node by sending the RSSI value which it computes with the receipt of a message from MN(mobile node). The parent node performs the task of finding new parent and sending of other control messages thereby saving the energy of a mobile node. EMA-RPL surely outperforms both RPL and ME-RPL but it also has some disadvantages. The parent node selected using this protocol might be for short duration hence more number of control messages are sent and less route stability occurs.

The authors in [18] presents a route stable strategy called Kalman positioning RPL (KP-RPL). In this, like in the other proactive approach mobile node can never be a parent node. All mobile nodes create a list called blacklist before finding routing information, which contains nodes that will be possibly unreachable to avoid any disconnection in the network. In this approach, a stable route is generated which improves the reliability of the network. However, there is an assumption in this protocol that mobility is low. This assumption makes this protocol limited. Besides this limitation, all the calculation are done by the mobile node itself so it consumes the resources of the mobile node. There is also inaccuracy in position calculation of mobile node that effect the route stability of the network and hence sometimes topology change of the network is frequent.

The author in [19] Extended Kalman Filter for mobile RPL (EKF-MRPL), uses an extended Kalman filter to determine the new position of the mobile node. As in EMA-RPL, the parent node notify about the mobility using RSSI value in this also. The mobile node based on its own coordinates and direction of motion calculates new parent. The coordinates are enhanced by Extended Kalman Filter. This protocol produces routes which are more stable and the number of control message sent are reduced by considerable numbers. Though it has some merits it also has some demerits. Since all the calculation are performed by a mobile node, the mobile node losses its energy faster. Moreover, it can also lead to an increased end-to-end delay.

III ROUTE STABLE EMA-RPL

The main aim of this protocol is to enhance the route stability in EMA-RPL with minimized use of resources of the mobile node. EMA-RPL is a proactive protocol that allows mobility to be detected before disconnection of network and hence resolving the issue of topology change in RPL. This work reduces the number of preferred parents changed in EMA-RPL. The energy dissipation of the mobile node is kept minimum as most of the calculations are done by a static node. The handover delay might increase depending upon the movement of the mobile node. The mechanism used for finding a stable route depends upon the speed of the mobile node with which it is moving and accordingly is the handover delay. The proposed protocol gives increased route stability with mobility enhanced and minimal data loss and handover delay

A. Overview and modification in EMA-RPL

EMA-RPL is an enhancement over RPL protocol which integrate mobility in RPL with minimum data loss and handover delay. In this protocol only static nodes can have child node, all the mobile nodes are leaf node. Same will be followed in the proposed work, all the mobile node will be leaf nodes to allow minimal data and path loss. The EMA-RPL is broadly divided into three phases. The first phase is

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:21:15 UTC from IEEE Xplore. Restrictions apply.

the detection and transmission phase. In this phase mobility is detected by associated node also called preferred parent using reduced RSSI (Received Signal Strength Indication) value. The detection of mobility in Route stable EMA-RPL follows the same rules as EMA-RPL. The next phase in EMA-RPL is reaction and prediction phase in which the next parent node is found. For this phase in this proposed protocol, there are some changes in control messages of ICMP. The last phase of EMA-RPL is notification phase in which a mobile node is notified about the parent node. The notification in this new protocol is modified according to the changes in the second phase.

B.Speed calculation of the mobile node.

The route stability in the presented algorithm is achieved using the speed of the mobile node. The speed is calculated using the change in the co-ordination of the mobile node. If it is considered that the initial coordinates of a mobile node are (x_i, y_i) and after Δt time, coordinates of the mobile nodes are (x_m, y_m)

Speed along x-axis $(v_x) = (x_m - x_i) / \Delta t$ (1)

Speed along y-axis $(v_y) = (y_m - y_i) / \Delta t$ (2)

Scalar quantity of speed $(v) = \sqrt{v_x^2 + v_y^2}$ (3)

C. Changes in ICMP Control messages

To incorporate route stability in EMA-RPL some changes are done in *ICMPv6* control message including changes for EMA-RPL. Beside modification for EMA-RPL like including of RSSI value and flag, some more fields are added in the ICMPv6 control messages (DIS, DIO) such as the coordinates of the mobile node.

D.The proposed Algorithm

After the formation of the DODAG tree for communication, mobility is detected using the reduced RSSI value. Once the mobility is detected, the parent node of a mobile node, which is also called an associated node in EMA-RPL protocol, starts detecting another suitable parent. The algorithm enters the next phase starts, which is searching for another parent node for the mobile node. Figure 1 gives the parent selection process of both the method. (a) shows initial condition while in (b) mobility starts and (c) tell which parent is selected

1.New parent node identification

The associated node broadcast the DIS message which includes the RSSI value and mobile node ID with *flag* set to 1. The time at which it broadcast the DIS message it also records the time and coordinates of a mobile node for speed calculation. Suppose the time at which DIS is broadcasted is noted down as t_0. Beside this t_0, a timer is triggered by an associated node to wait to receive reply of DIS message broadcasted. The broadcasted DIS message is received by mobile node and other neighboring static nodes. As soon as mobile node (MN) receives DIS message from associated node, it on the basis of RSSI value in the DIS message decides whether it will continue sending data or will stop to avoid any data loss. The MN also broadcast DIS message with *flag* =2 in order to be detected by static nodes (SN). As the broadcasted message is received by all neighboring SN and associated node, the associated node comes to know the current co-ordinates of the mobile node. As soon as the broadcasted message is received by associated node it note down the time as t_n.

The associated node calculates the speed of mobile node as follows-

It is considered that initial co-ordinates of mobile node were (x_i, y_i) and new co-ordinates are (x_m, y_m). The change in time is t_n-t_i. Then the speed is given as

Speed of mobile node = $\sqrt{\left(\frac{x_m - x_i}{t_n - t_i}\right)^2 + \left(\frac{y_m - y_i}{t_n - t_i}\right)^2}$ (4)

All the SN who received DIS message broadcasted by mobile node sends a unicast DIO message which includes RSSI value calculated using the mobile node's DIS message to the AN. The associated node on the basis of received RSSI values compares them and find the best RSSI value. The node with best RSSI value is considered for next phase.

2.Stable route selection

The node which is selected with best RSSI value is notified by the Associated node that it will act as associated node with some defined conditions. The notification message also includes the speed with which mobile node is moving and time span (t_n-t_0) under which speed was calculated. The AN also notify the mobile node about new associated node. The new associated node will then send a DIO message with *flag=3* to the mobile and informing that it will just act as a temporary parent to find a stable associated node. Upon receiving the DIO message with *flag=3,* the mobile node sends a unicast DIO message to new associated node providing its co-ordinates and starts a timer.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:21:15 UTC from IEEE Xplore. Restrictions apply.

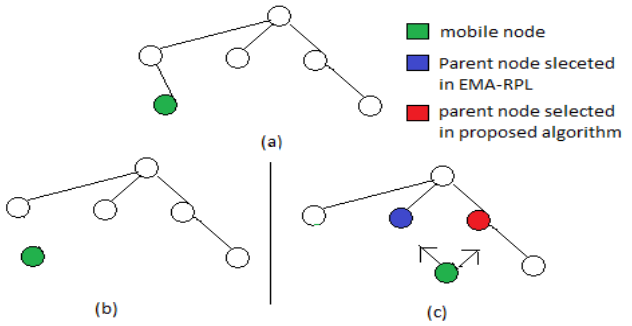


Figure 1: Difference between parent selection in EMA-RPL and proposed algorithm

As soon as the new associated node receives DIO message from the mobile node it records the co-ordinate of MN and time t_{a_0} at which the message is received. When the timer expires at the mobile node, it again sends a message with its coordinates. The new AN note down the time and calculates the speed of the mobile node as was calculated by old AN. According to the new speed s_n calculated it decides whether it will be selected as a parent node or not. Following conditions may occur

- If the new speed s_n calculated is greater than or equal to old speed s_i. Then it is definite that the mobile node is still in motion and the selected new temporary associated node will not give a stable root.

- The new speed is in between $0.3 * s_i$ to s_i , that is in between 30% of initial speed and the initial speed then there are chances that new associated temporary node will not give a stable route. It depends upon the speed whether it will give a stable route or not.
- If the calculated new speed s_n is less than $0.3 * s_i$ then there is a high probability that temporarily associated node may give a stable route.

According to the above conditions following an event may occur

- If the speed satisfies first two conditions then if a new temporary associated node is made as a preferred parent the path would not be stable. Hence instead of making the new node as a preferred parent the new associated node act as a temporary parent node and find another more reliable preferred parent using the above algorithm.

- If the speed of the mobile node satisfies the third condition then temporarily associated node is selected as preferred parent. It notifies the mobile node about this. The mobile node upon receiving this notification DIO message updates its parameter like the preferred parent, route to the root node etc. and notify about this to the temporarily associated node.

The route stability can be defined as the inverse of the rate of change of preferred parent. So if for a mobile node's number of preferred parent change or total number of preferred parent selected in a time period t is p, then route stability is

Route Stability $R_s = \frac{t}{p}$ (5)

A higher value of the rate of change of preferred parent gives less stable route while a lower value provide route with more stability. It can also be said that route stability is inversely proportional to the rate of change of the parent node of a mobile node. For calculation, the average of all mobile node route stability is taken.

E. Correctness of speed calculated at different time.

As in the calculation of speed, the delay associated in the message sending have not been considered hence it might seem that the calculated speed is not accurate. There are two cases in that as follows-

Case I: When the propagation delay is almost same.

Initially when parent node came to know about the motion of mobile node and received reduced RSSI value and co-ordinates it note down the time to calculate speed. Suppose the time to reach a message from mobile node to associated node is ∂t_i . Then the time at which the mobile node was at the co-ordinates given in message is

The actual time at which mobile node was on co-ordinates $(x_i, y_i) = t_i - \partial t_i$ (6)

Similar at later if ∂t_n is the delay then actual time is

The actual time at which mobile node was on co-ordinates $(x_m, y_m) = t_n - \partial t_n$ (7)

The difference in time $\Delta t = (t_n - \partial t_n)$- $(t_i - \partial t_i)$ (8)

Now if a delay is approximately equal then

The difference in time is $\Delta t = t_n$-t_i(approximately) (9)

As ∂t_n and ∂t_i gets cancelled as they are equal.

Hence the calculated speed is accurate.

Case II: When the propagation delay difference is considerable.

If the propagation delay is different in two messages from a mobile node to the old parent node, then some difference will

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:21:15 UTC from IEEE Xplore. Restrictions apply.

also be there when the new temporary associated node calculates the speed. Since this difference exists in both condition hence the speed calculated at two different times will have the approximate same error. Due to error, the speed calculated will be a little less than the actual speed, but this slow speed is recorder at the second time also, hence it will not make much difference and the algorithm works fine with this error in speed also.

IV. PERFORMANCE EVALUATION

In this section, the performance evaluation of the proposed Route Stable EMA-RPL protocol through simulation performed is discussed. Most of the comparison is done between the proposed algorithm and EMA-RPL with some comparison with other protocol

A. Simulation setup

An open-source operating system for IoT called Contiki IPv6/

6lowpan platform is used and the open-source implementation of RPL named ContikiRPL that is widely used [20]. The Contiki OS provide Cooja simulator which simulates a different type of Contiki nodes. For the proposed method Contiki is used as it provides built-in RPL and contains many types of motes which simulate real-life WSN very closely.

Most of the comparisons are done between the proposed protocol and EMA-RPL protocol. Three parameters are considered for comparison as handover delay, Route stability and energy consumption. For simulation following parameters are used in cooja setup

Simulation parameters

The simulation of the proposed method was done with the following parameters set in cooja simulator. Beside those skies, motes were used with

Table 1. Parameters for simulation

Data rate	30 packets/s
Transmission range	50m
Threshold RSSI	-90dBm
Simulation time	800s
Sensing Area	150*150 m^2
Transmitting power	0.23 mW

B. Route Stability

The comparison of EMA-RPL and proposed route stable EMA-RPL is shown in fig 2.

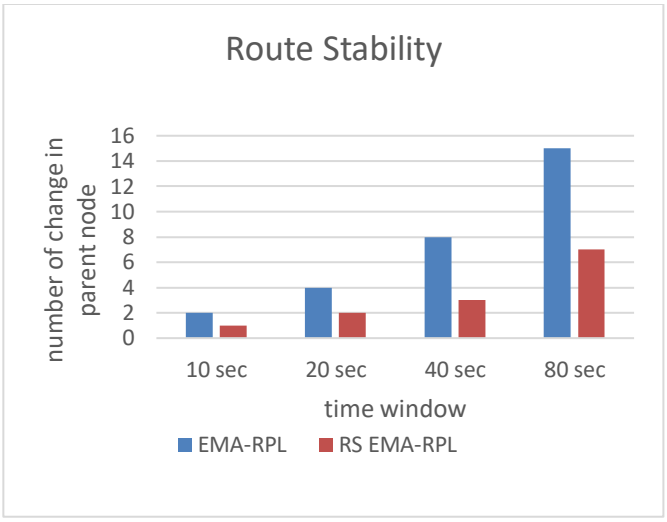


Figure 2: Simulation results of the Route stability

It is calculated as many parent change in a given time window. The route stability in EKF and the proposed algorithm is almost the same hence our method is compared with EMA-RPL only. Even in many cases, the proposed algorithm outperforms EKF by a little margin but here our main focus is comparing it with EMA-RPL. Route stability is inversely proportional to the rate of change of parent node as

Route stability $R_s =$ 1/rate of change of parent node (10)

C. Handover delay

The handover delay is defined as the difference between the time when the last message was sent by the mobile node before disconnection and the time at which the first message is sent after the new connection is established. Though EMA-RPL outperforms our proposed algorithm in handover delay as shown fig 3 the difference is not as huge as RPL. It might seem that handover delay is considerable but it is still comparable to many mobilities enhanced protocol. The handover delay depends upon the rate of parent node change. As when the rate of change of parent node is high, handover delay is less as after disconnection a new connection is found fast enough compared to the case in which rate of change of parent is low while condition like the speed of mobile node, mobility and other things are same in both case

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:21:15 UTC from IEEE Xplore. Restrictions apply.

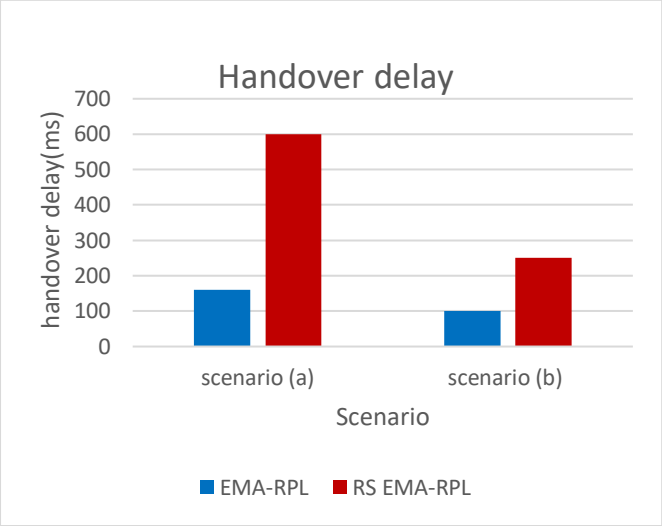


Figure 3: Experimental results of the average handover delay

D. Energy Consumption.

The energy consumption of a mobile node in the proposed algorithm and EMA-RPL is same. Hence here EKF is compared and proposed algo as both algorithms produce the route stable path. As it can be seen be from the graph in Fig 4 that energy consumption of a mobile node in EKF algo is more than that of Route stable EMA-RPL. The energy consumption of a static node is almost the same in both algorithms.

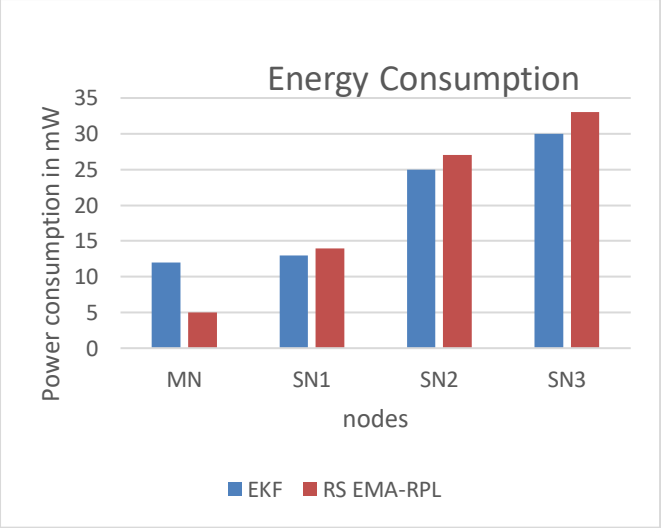


Figure 4: Average consumed energy after successfully simulating protocol for a considerable time

V. CONCLUSION

A new algorithm is proposed to increase the route stability in EMA-RPL protocol with minimum data loss and least topology change. This approach is proactive as mobile node before disconnection with the previous parent finds a new parent.

To make minimum utilization of mobile node resources, most of the steps and calculation are done by the parent node. Though the proposed protocol might have given more accurate results if the speed calculations were performed by the mobile node itself it may cost resources of a mobile node which need more attention than accuracy. Though handover delay has increased but comparing to RPL protocol it is still less. As future work, the handover delay reduction can be studied and better result can be obtained.

REFERENCES

[1] P.Lopez, D.Fernndez, A.J. Jara and A.F. Skarmeta, "Survey of Internet of Things Technologies for Clinical Environments", AINA Workshops, 2013, pp. 1349-1354.

[2] A.J. Jara, L. Ladid and A. Skarmeta, "The Internet of Everything through IPv6: An Analysis of Challenges, Solutions and Opportunities", Journal of Wireless Mobile Networks (UCDA), vol. 4, pp. 97-118, 2013.

[3] J. Yick, B. Mukherjee and D. Ghosal, "Wireless sensor network survey", International Journal of Computer and Telecommunications Networking, vol. 52, pp. 2292-2330, 2008

[4] P. M. M. Carlos Abreu and Manuel Ricardo, "Energy-aware routing for biomedical wireless sensor networks," Journal of Network and Computer Applications., vol. 40, pp. 270–278, 2014.

[5] D. Culler, D. Estrin and M. Srivastava, "Overview of Sensor Networks", IEEE journal of Computer Society, vol. 37, pp 41-49, Aug. 2004.

[6] T. Winter, P. Thubert, A. Brandt and J. Hui Rpl: IPv6 Routing Protocol for Low-Power and Lossy Networks, Internet Proposed Standard RFC 6550, 2012.

[7] Z. Shelby, C. Bormann, "6LoWPAN: The Wireless Embedded Internet", book, Internet Engineering Task Force (6LoWPAN-WG), 2009.

[8] G. Montenegro, N. Kushalnagar, J. Hui and D. Culler, "Transmission of IPv6 Packets over IEEE 802.15.4 Networks", IETF RFC 4944, Sept. 2007.

[9] Md.S. Hossen, A.F.M.S. Kabir, R.H.Khan and A. Azfar, "Interconnection between 802.15.4 Devices and IPv6: Implications and Existing Approaches", IJCSI, vol.7, Jan 2010.

[10] T. Winter, P. Thubert, A. Brandt, T. Clausen, J. Hui et al. "RPL : IPv6 Routing Protocol for Low power and Lossy Networks", RFC 6550, Mar. 2012.

[11] L.Bartolozzi, F.Chiti, R.Fantacci, T.Pecorella and F.Sgri lli, "Supporting monitoring applications with mobile Wireless Sensor Networks: The eN Route forwarding approach", IEEE International Conference on Communications , pp. 5403-5407, June 2012.

[12] K. C. Lee, R. Sudhaakar, J. Ning, L. Dai, S. Addepalli, "A Comprehensive Evaluation of RPL under Mobility", Hindawi Publishing Corporation, International Journal of Vehicular Technology Vol. 2012.

[13] O. Gaddour and A. Koubaa, "RPL in a nutshell: A survey", ELSEVIER Journal of Computer Networks, pp. 3163–3178, 2012.

[14] I. El Korbi, M. B. Brahim, C. Adjihy and L. A. Saidane, "Mobility Enhanced RPL for Wireless Sensor Networks", IEEE Third International Conference on the Network of the Future (NOF), pp. 21–23, Nov. 2012.

[15] M. Bouaziz and A. Rachedi, "A survey on mobility management protocols in Wireless Sensor Networks based on 6LoWPAN technology", Elsevier Journal of Computer Communications, vol. 74, pp. 3–15, 2016.

[16] M. Bouaziz, A. Rachedi, A. Belghith,M. Berbineau and S. Al-Ahmadi, " EMA-RPL: Energy and mobility aware routing for the Internet of Mobile Things", Future Genereration Computer System, vol.97, pp. 247–258, Feb. 2019.

[17] N. Patwari, J.N. Ash, S. Kyperountas, A.O. Hero, RL Moses and N. Correal, "Locating the nodes. Cooperative localization in wireless sensor networks", Signal Process Magazine, vol.22, pp. 54-69, Aug. 2005.

[18] A. Ribeiro, I. D. Schizas, S. Roumeliotis, and G. B. Giannakis, "Kalman filtering in wireless sensor networks," IEEE Control System., vol. 30, no. 2, pp. 66–86, Apr. 2010.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:21:15 UTC from IEEE Xplore. Restrictions apply.

[19] M. Bouaziz, A. Rachedi and A. Belghith, "EKF-MRPL: Advanced mobility support routing protocol for internet of mobile things: Movement prediction approach.", Future Genereration Computer Syststem, Dec. 2017.

[20] W. Dong, C. Chen, X. Liu, and J. Bu, "Providing OS support for wireless sensor networks: challenges and approaches", IEEE Communications Surveys & Tutorials, vol. 12(4), pp. 519-530, 2010.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:21:15 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Second International Conference on Inventive Research in Computing Applications (ICIRCA-2020)
IEEE Xplore Part Number: CFP20N67-ART; ISBN: 978-1-7281-5374-2

A Survey of Application of ML and Data Mining Techniques for Smart Irrigation System

Mr. Himanshu Nandanwar
Information Technology
Delhi Technological University, Delhi, INDIA
himanshunandanwar9cm0@gmail.com

Mrs. Anamika Chauhan
Information Technology
Delhi Technological University, Delhi, INDIA
anamika@dce.ac.in

Mr.Dushyant Pahl
Information Technology
Delhi Technological University, Delhi, INDIA
dushyantpahl@gmail.com

Mr.Harshit Meena
Information Technology
Delhi Technological University, Delhi, INDIA
meenaharshit30@gmail.com

Abstract. This paper reviews our current research in agriculture analytics on an open-source platform using data mining and machine learning techniques. Various sensors are used to collect data that provides real-time analytics on the weather forecast, soil moisture, air temperature, PH, humidity. The smart irrigation system is paired with different hardware and development application. The science of machine learning and data mining plays a significant role here. It is a data analysis process using multiple models and algorithms to trained data directly. Machine learning and data mining to hold the irrigation system in r leaner manner. For instance, it optimizes the use of water and provides a critical amount of water and fertility to increase production efficiency, reduce manpower involvement, and reduce crop diseases. The survey tests the effect of applicable techniques and how these techniques boost efficiency.

Keywords - Data mining, machine learning, smart irrigation system, real-time data analysis, sensor data.

I. INTRODUCTION

As it is already known that, agriculture is the largest and significant water consumer. The irrigation system consumes around 65-70% of global water, while other industrial and domestic sectors consume 20% and 10-15% of global water respectively [1]. This consumption of water diversify country to country, region to region and have different water consumption rates. Periodically the population is increasing, so the demand and supply of food are also increasing exponentially. There are many more problems and factor which affects the irrigation system like whimsical and unpredictable climate condition which impact in the crop growth, inadequate rainfall, scarce and infrequent natural water resource availability, lack of knowledge, and obtuseness regarding new technologies that can lead to the higher productivity using minimal resources. Subsequently and eventually aware of this problem, agriculture is the main source of employment especially in India and it contributes to 17-20% of GDP [2]. Crops require water for proper growth. Traditional methods of irrigation required a large sufficient amount of water resources and manual intervention. Over irrigation and under irrigation will affect the soil moisture, increase in crop diseases, and will resource efficiently and required almost negligible manpower. A smart automated irrigation system uses various technologies such as big data, IoT, machine learning, wireless sensor networks, cloud computing. This automated irrigation system is entirely streamlined, systematic, labour-saving, energy-efficient, well regulated, and knowledge dependent. Researchers are closely linked to production and every other part of the recent development and cleverness of agriculture aims to improve crops productivity, reduce and prepare for plant conditions, boost modern mechanisms and automated agriculture and agro-industry [3]. The machine learning and data mining approach can be summarized as data acquisition, data pre-processing, neural networking, model testing & forecasting, etc. Internet of things (IoT) will supply data set of high quality. Such data significantly improve the use and accuracy of the resulting devices for machines in agriculture. Training and learning algorithms can improve precision, in particular for applications requiring high precision. This paper proposed the data mining and machine learning techniques. Data mining is used for understanding, data preparation, data transformation, and data modelling. Data mining refers to the extraction of knowledge from large data sets. Data mining is about identifying the hidden, legit, and potentially useful patterns. Data mining is connected to the data set and data warehouse. The goal of data mining is to make secure and easy access to information to the user, help to provide information from the data sets. It helps in taking a quick and effective decision. This paper uses KNN (k nearest neighbour) techniques for smart automated monitoring and controlling irrigation systems for decision making. This proposed system will use machine learning algorithms for water management, crop productivity and soil management, detection of soil diseases and weeds, and for quality of crops.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:19:52 UTC from IEEE Xplore. Restrictions apply.

Smart irrigation system answers 3 problems of agriculture: -

• Water requirements for crop/yield (on regular basis)
• Use of fertilizer for the crop at a specific and certain time based on
micronutrients (boron, chlorine, iron, zinc, manganese) and
macronutrients (nitrogen, phosphorous, potassium, sulphur,
calcium)
• Use of pesticides for the crop at a specific and certain time based on
environmental conditions such as temperature, humidity, pressure,
pollution, etc.

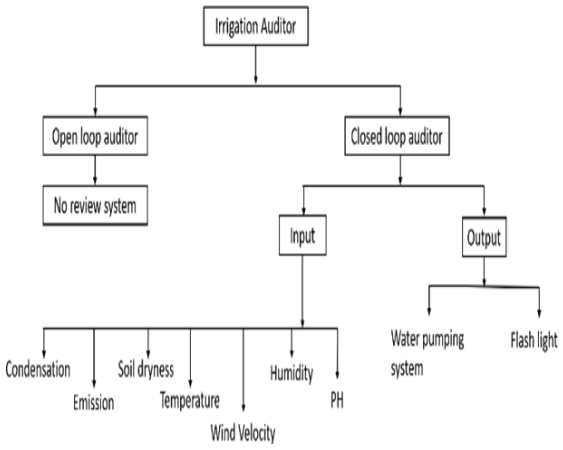


Fig.1. The functioning of the smart Irrigation System.

In a smart irrigation system, the water requirements depend on environmental factors like temperature, humidity, ground slope, the intensity of irrigation, probability of rainfall. Water required in the irrigation module Hector named. Irrigation depends on certain variables like these Crops, Eto(evaporation), crop stage, soil condition, Rate), temperature, humidity, etc. present-day, and the following day. Diagram reveals how the intelligent irrigation system functions Take temperature, name of crop, crop level, soil condition, cultivation as an input element. All these inputs are then pre-processed as water per hectare is required, result. Block diagram of the functioning of the smart irrigation system shown in Fig1.

II. LITERATURE REVIEW

A. IRRIGATION AUDITOR

Irrigation is a method of controlling the amount of water to crop at required intervals [4]. Efficient irrigation helps in increasing crop productivity. It will protect from the famine. Good irrigation also helps in maintaining water condition in soil and increase the water content of the crop and dissolve nutrients and makes them available to the crops. Irrigation auditor is large groups into two types: - open-loop auditor and closed-loop auditor. Classification of irrigation model auditor is shown in Fig.2. Open-loop auditor means there is a pre-defined state such as an interval of irrigation, frequency of irrigation, irrigation session, and amount of water required for irrigation. The closed-loop auditor is a feed-forward network system that controlled the. The system automatically will conclude the result based on balancing between the pre-processing data and concurrent computation.

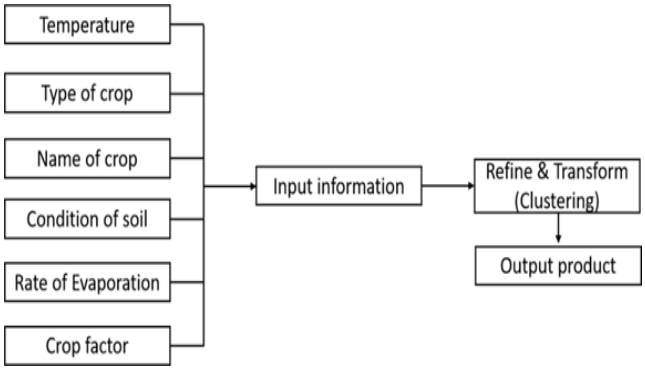


FIG.2. Classification of Irrigation Model Auditor.

B. SENSOR DEVICE.

To measure the different parameters of crops, different sensors are required for measuring temperature, air pressure, weather conditions, humidity, leaf sensing, soil moisture of crops [5]. The real-time value will be estimated using various sensors deployed in the irrigation field. Sensors should be low cost, effective, and accurate for measuring the data. Climate based controller also known as evapotranspiration (ET), it uses the native and localized weather data to regulate and re-arrange irrigation schedule day to day plan. Evapotranspiration controller help in reducing water usage. Soil moisture sensor controller senses the soil moisture content of the crop from the root zone. It helps in maintaining the volumetric water content of the soil. Rain and freeze sensor is used to stop unnecessary irrigation during the rainy season and assist in saving the water resource. Water management especially saving water depends on the rainfall. Wind sensors help in identifying the average wind speed in nearby areas of crop. Watering during windy conditions reduce irrigation uniformity and decrease the amount of water infiltrating into the soil. Various types of generic controller and sensors are shown

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:19:52 UTC from IEEE Xplore. Restrictions apply.

on Fig.3.

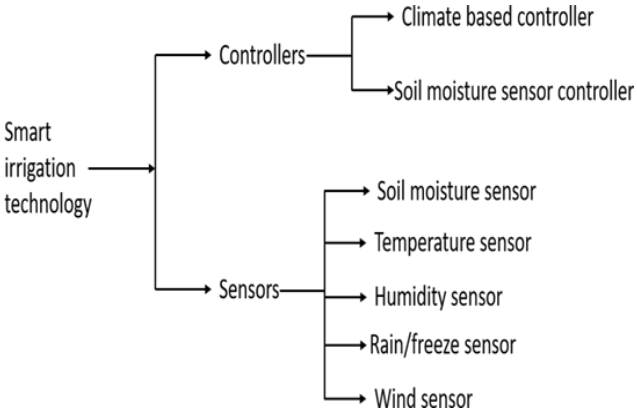


FIG.3. Type of Generic Controllers and Sensors.

C. GENERIC MODEL.

There are many traditional methods of irrigation. Generally, farmers surface irrigation, lever Irrigation system for irrigating their fields. These methods require more amount of water and it's time-consuming and laborious [6]. Due to traditional methods of irrigation wastage of water is there and there is a possibility of overwatering band wasteful runoff. So different smart irrigation system is available based on hardware/software as well as mobile-based irrigation system were introduced. A smart irrigation system uses various hardware (sensor) and software (desktop application) components. This smart system collects real-time data and helps in extracting knowledge, useful patterns from it. Some system has complex software which provides information from various sources such as weather station, sensor real-time data, data from the internet, data from different database which are useful for making a quick and smart decision. Multiple types of sensors are used to sense the environmental conditions of crops yield such as temperature, humidity, PH, soil moisture, air temperature, soil temperature, radiation, wind speed, etc. These measured values help in monitoring and controlling the crop productivity, amount of water required, amount of fertilizer, and pesticides needed. This software uses various techniques and algorithms such as system, artificial intelligence method, data mining, Machine learning techniques for data interpretation, and processing.

Their mobile-based irrigation system also uses Arduino UNO. Arduino UNO is the brain and heart of the system. It consists of the microcontroller which controls the water pump ON/OFF system. The soil moisture sensor is connected to the microcontroller, it senses the soil moisture and sends the measurement to the Arduino. If the value of soil moisture is below the eye threshold value, it sends the SMS to switch ON the water Motor. If the value is above the threshold, it sends SMS to switch OFF water motor through GSM. If at some point in time, sensors stop working it will not provide accurate data & results and this will affect the performance of the system. Using data mining and machine learning on the smart irrigation system, the required amount of water pattern is observed and calculated, where it's a correlation with the device easily and accurately. Fertilizer system fertilizer requirement varies with climate condition, irrigation method, and type of soil used for irrigation. Soil nutrient value, the amount of soil organic matter, crop management help in making fertilization decisions [7] by calculating information from various parameters.

This fertilizer system is generally work in 4 phases: -

- System requirements: - understanding of soil, it's nutrients, climate condition, type of fertilizer & its properties, type of crop.
- Fertilizer capability: - type of fertilizer, knowledge of last fertilizer used, fertilizer rate of absorption.
- Association between system and fertilizer.
- Correlation of fertilizer with different crops.

Structure and techniques of pesticides -Pesticides are used to protect the crop against various insect pests, weeds and fungal diseases. It prevents contamination of foods by rats, mice, flies and other insects. It protects human health by stopping the fungi from reaching food crops [8]. Pesticides allow farmers to produce safe, high-quality food at affordable prices and help farmers grow more food crops inside limited land. Comparison of different machine learning and data mining techniques presented in table 4.

D. Deductive process of learning in machine learning: -

Machine learning plays a major role in the intelligent irrigation system based on different technologies.

It works on the computational and statistical techniques to learn data collected from the agricultural field. The process of learning concentrates on learning from training data to achieve a specific task. Machine learning [9] generally categorized into two learning techniques supervised learning and unsupervised learning. Classification of a different machine learning algorithm in the irrigation system is shown in Fig.4. The irrigation system uses the learning process as such. The most common learning methods are supervised learning which

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:19:52 UTC from IEEE Xplore. Restrictions apply.

uses labelled data used for training purposes as input data. It is composed of both input factor and output requirements. Unlabeled data used in unsupervised learning for training purposes. The clustering of training data into a common category is performed based on the related existence of the variables. The following sections address servals irrigation techniques that have been implemented using ML.

The neural network method is commonly employed using ML in the irrigation system. It has different forms and various definitions are used. In the input layer, multiple parameters are given. For better outcomes, the input layer includes more than one node or neuron. It may be the growth stages of the plants or different forms of environmental conditions. The hidden layer is then placed after the input layer. It accesses the data from every node in the input layer. The output layer is then mounted behind the hidden layer.
That shows the desired output or planned output in the hidden layer according to trained data. Serval algorithms are proposed for irrigation system, each algorithm with its different parameters is observed [10] [11] [12] and functionalities are discussed in table 2&3.

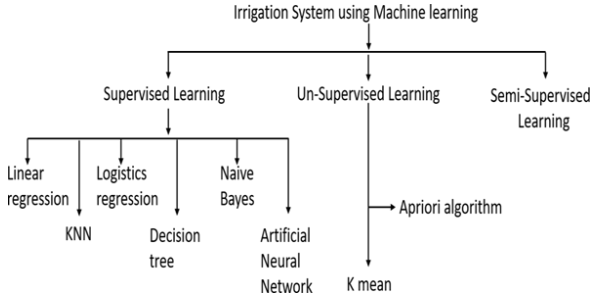


FIG.4. Classification of Different Machine Learning Algorithms in Irrigation System.

Table.1. Management of Soil Using Machine Learning.

S. NO.	Specification	Perceive Feature	Performance	Technique
1	Soil Moisture	Speculative nature of force on a no till chisel	Evaluation of soil moisture	MLP & RBF
2	Soil Temperature	Temperature of air, radiation of solar and barometric pressure	Evaluation of soil temperature	ECM & SaE-ELM
3	State of soil	Optimum layer of soil-soil sample(140) in cultivated field	Calculating the soil, organic carbon, soil moisture content and amount of nitrogen	LS-SVM & Cubist
4	Classification in soil	Accessible evaporation transpiration data and condensation	Checking out the soil process its dryness and wetness in binary grade.	KNN, Back propagation

Table.2. Management of Water System using Machine Learning.

S. NO.	Specification	Perceive Feature	Performance	Technique
1	Saturation point temperature	Temperature of air, barometric pressure, solar radiation, soil moisture evaporation	Forecast temperature at e saturation point(daily)	ELM & ANN
2	Evapotranspiration (ET)	Temperature of air, radiation emits by natural resource, evaporation	Forecast of evaporation rate depends on weather station(weekly)	ELN & ANN
3	Evapo transpiration	Temperature of air, barometric pressure,speed of wind, concentration	Evaluation of evaporation rate for local and pooled point	ANN, RNN , ELM & G

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:19:52 UTC from IEEE Xplore. Restrictions apply.

E. Machine Learning and Data mining techniques in smart irrigation system: -

Methods provide a learning method to understand "experience" to perform a task (training data). The data is established by ML examples. Examples. A group of attributes also known as Characteristics of variables are usually represented as individuals Exhibit. An enumerated function is nominal. It could be binary (i.e., 0 or 1), or numerical (i.e., + or B-). In a particular way, the task is to measure the performance of the ML model Improved metric performance over time. Various statistics and math modelling are used for performance calculation ML and algorithms design. After the process of learning The trained model can use the acquired experience Classify, predict and close new during training Examples (test data) Examples. ML-work relies on the signal for learning the method of learning divided into two Unregulated or unregulated learning groups. Data is shown Inputs and correct outputs with example regulated learning and the aim is to establish an overall framework Rule outputs for maps. Just partial inputs can be rendered in some cases Get with missing or only given target outputs Competence feedback in a complex environment (Learning in enhancement). The know-how (trained model) used to predict the lack of test data in the monitored setting The outputs (labels) are available. There is, however, no distinction between Sets of training and testing when unlabeled data Uncontrolled preparation. Uncontrolled preparation. The student analyzes the information input to recognize secret designs.

Data mining [13] is a technique to extract useful information from the sea of data. It identifies any hidden patterns, future trends, and behaviours and allowing businesses to make decisions accordingly.

The data mining technique can be corporated with the cloud computing system. The system provides services such as field monitoring, market-based analysis, storing of data, decision-making process for the farmers. Cloud is commonplace for storing all types of data. The system can support the agriculture sector because it provides features such as minimizing the spending, minimizing the labour, time reduction, pest control, and at last manages the farmer's marketing and banking information. Data mining is a tool for collecting valuable data from the oceans. It identifies any hidden patterns, future tendencies, and conducts and enables companies to decide accordingly. Various methodology for research in data mining is shown in table.3.

With the cloud computing system, data mining technology
can be operated. The system offers services such as field surveillance, market-based analysis, data storage, farmers
decision-making. The cloud is used to store all data types.
This program will benefit agriculture as it offers features such as a minimization of investment, minimization of energy, reduction of time, and control of pests.

Table.3. Methodology for research in data mining

Data chain stages	State of art	Key problems
Data collection and review	Detector, Open information	Data visualization
Stocking of data	Framework based on cloud	Fast and secure data access
Data Collection	Standardization	Diversity of information sources, Cleansing of information, Automation of planning
Analysis of data	Ontology of the decision, Computing cognitive	Heterogeneous Semantic, Analysis of real-time data, Scalable
Reporting of information	Viewing of data	Property, Confidentiality, New model of business

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:19:52 UTC from IEEE Xplore. Restrictions apply.

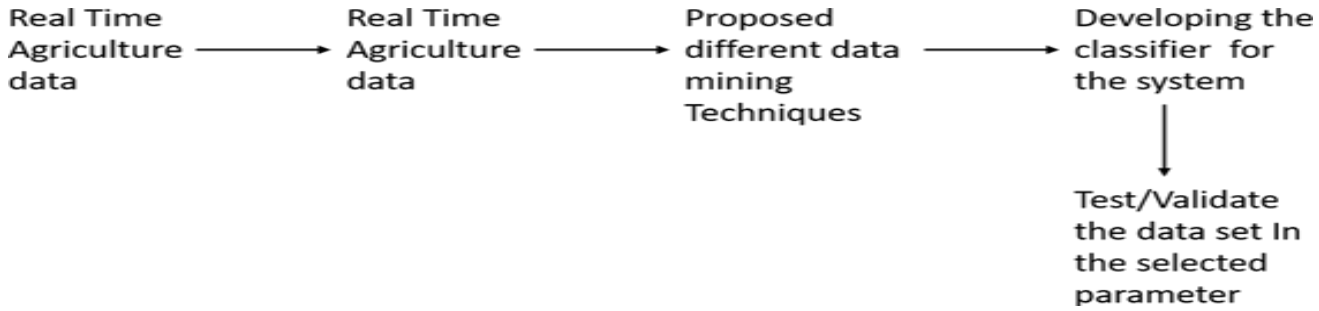


FIG.5. Flow Diagram for proposed Model

III. CONCLUSION

This paper has implemented the smart farming system approaches which will result in full crop yield. This method is very popular in terms of the outcomes as its success relative to the methods built before. The clustering algorithms used in this method are the pattern matching that elegantly returns the perfect output according to the p input parameters. This system does not require any potential interventions or time-to-time handling or modification of the data because of its autonomous existence. It will specifically concentrate on crop growth & production, and by applying its strategies it will increase productivity. The smart farming system will also warn farmers about critical weather conditions that will make any imaginable anomaly sustainable once again. Water is a resource-limited resource and its conservation is nowadays the biggest problem, but using this method would help in the proper use of water and no excess or under-supply. To summarize all the method of smart farming is an idol solution for advance irrigation.

According to this study, it enhances irrigation systems by applying machine learning in agriculture. It helps make efficient use of the water and reduces water wastage. This method of incorporation of automated data processing, data recording, and decision implementing machine learning is a complete ration based framework. It increases the degree and quality of output. This survey can be useful in providing farmers with prior knowledge to implement machine learning techniques with irrigation action system based on their needs and increase productivity.

IV. FUTURE WORK

Many potential works can be expanded from the proposed concept, and a large number of characteristics reflect the general view of the plant. The orientation of the leaves, the stains in the leaves, the Galactose trees, and their effects on appearance and many other patterns help to identify plant problems. The machine learning and data mining techniques and their principles should also be implemented with several implementations which make the system smart and in the future with the effective combination of these techniques, the efficiency of the system can be boosted. The proposed model's flow diagram is discussed in Fig.5.

Table. 4. Comparison of various existing machine learning and data mining techniques

Algorithm used	Technique	Significance	Data processing & transformation
Extreme Learning Machine(ELM) & kernel based EML [14]	If different between backpropagation, SVM & ELMAN simulation techniques.	To track greenhouse temperature & humidity.	WSN based on ZigBee

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:19:52 UTC from IEEE Xplore. Restrictions apply.

Fuzzy logic [15]	It used to forecast climate using humidity & temperature. It gives information about soil moisture from the sensor and helps in crop production.	Predict the type of various crop to be grown using data from the soil pH, temperature & humidity sensor. Related field Use GSM to forward information to the mobile former.	WSN based on GPRS and GSM.
Data mining approach apriori algorithm, linear regression [16]	Using the association rule between various parameters to find a relationship (Temperate, humidity, pH). Using linear regression on input help to build the relationship among different parameter.	Analysis of sensor data to predict appropriate crop temperature, humidity & soil moisture.	Based on WSN
Naïve Bayes algorithm [17]	Computation of soil moisture, nutrients, amount of water required amount of fertilizer needed using real-time data from the sensor.	To accurately predict detect test/ disease using real-time sensor data to control & moisture the crop quality.	WSN Based on ZigBee
Bayesian statistics [18]	From the data given by sensor and another resource, is used to understand the correlation between crop quality, image processing for weeds, insects.	A cloud-based The decision-making system uses the highest algorithm to optimize the activity of the sensor, schedule Irrigation based on a forecasting sensor.	Wireless Network & MQTT Raspberry Pi-based sensor & Arduino.
Machine vision and statistical analysis [19]	It uses digital image processing using canny detection technique and used to analyze the relationship b/w various parameter of environmental condition	To discover the top foremost condition to grow & increase the crop productive	Environmental surveillance & IOT based wireless imaging network Wi-Fi & ZigBee
SVM, linear regression and Neural Network [20]	This data mining method is used to identify the correlation between dataset and help in examine crop quality & production.	Supporting WSN-based device to provide an optimum growing condition for roses in a greenhouse	WSN based on ZigBee, sensor data is used to examine the temperature, humidity, soil moisture
RF, KNN, logistic regression & linear regression [21]	Using this algorithm, temperature, soil moisture, humidity, relative pressure, vapour pressure, micro & macronutrient of soil sensor data is identified. In all this technique, KNN has an efficient and effective technique.	For the collection & analysis of data to establish a predictive model that can be used to predict the plantation disease outbreak.	WSN based on ZigBee

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:19:52 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] Milos Brajovic, Stefan Vujovic, Slobodan Dukanovic, "An Overview of Smart Irrigation Software",4th Mediterranean Conference on Embedded Computing, MECO – 2015.

[2] Megersa G, Abdulahi J (2015) Irrigation system in Israel: A review. International journal of water resource and environment engineering 17(3): 29-37.

[3] Nanyang Zhu, Xu Liu, Ziqian Liu, Kai Hu, Yingkuan Wang, Jinglu Tan, Min Huang, Qibing Zhu, Xunsheng Ji, Yongnian Jiang, Ya Guo International Journal of Agricultural and Biological Engineering 11 (4), 32-44, 2018.

[4] Munoth P, Goyal R, Tiwari K (2016) Sensor based Irrigation System: A Review. International Journal of Engineering Research & Technology (IJERT) 4(23): 1-5.

[5] Maduranga, M., & Abeysekera, R. (2020). MACHINE LEARNING APPLICATIONS IN IOT BASED AGRICULTURE AND SMART FARMING: A REVIEW. *International Journal of Engineering Applied Sciences and Technology*, 24-27.

[6] Review - Machine Learning Techniques in Wireless Sensor Network Based Precision Agriculture Yemeserach Mekonnen-Srikanth Namuduri-Lamar Burton-Arif Sarwat-Shekhar Bhansali - Journal of The Electrochemical Society – 2020.

[7] https://content.ces.ncsu.edu/extension-gardener-handbook/1-soils-and-plant-nutrients

[8] https://www.britannica.com/science/plant-disease/Obtaining-disease-resistant-plants

[9] Coopersmith EJ, Minsker BS, Wenzel CE, Gilmore BJ (2014) Machine learning assessments of soil drying for agricultural planning. Computers and Electronics Agriculture 104: 93-104.

[10] Johann AL, de Araújo AG, Delalibera HC, Hirakawa AR (2016) Soil moisture modeling based on stochastic behavior of forces on a no-till chisel opener. Computers and Electronic Agriculture 121: 420-428.

[11] Feng Y, Peng Y, Cui N, Gong D, Zhang K (2017) Modeling reference evapotranspiration using extreme learning machine and generalized regression neural network only with temperature data. Computers Electronics and Agriculture 136: 71-78.

[12] Mohammadi K, Shamshirband S, Motamedi S, Petkovic D, Hashim R, et al. (2015) Extreme learning machine based prediction of daily dew point temperature. Computer and Electronics Agriculture 117: 214-225.

[13] file:///E:/02_synopsis.pdf

[14] Qi Liu, Dandan Jin, Jian Shen, Zhangjie Fu, and Nigel Linge, "A wsn-based prediction model of microclimate in a greenhouse using extreme learning approaches," In 2016 18th International Conference on Advanced Communication Technology (ICACT), pages 730–735. IEEE, 2016.

[15] Sangeetha Ravi, G Bhavani, Binu Ruby Sunny, P Abirami, and V. Divya, "Multidisciplinary effective prediction of crop using iot and wsn," International Journal of Advanced Research in Computer Science, 7(1) (2016).

[16] Jirapond Muangprathub, Nathaphon Boonnam, Siriwan Kajornkasirat, Narongsak Lekbangpong, Apirat Wanichsombat, and Pichetwut Nillaor, "Iot and agriculture data analysis for smart farm," Computers and Electronics in Agriculture, 156, 467, (2019).

[17] Hemantkumar Wani and Nilima Ashtankar, "An appropriate model predicting pest/diseases of crops using machine learning algorithms," In 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS), pages 1–4. IEEE, 2017

[18] Jetendra Joshi, Siddhanth Polepally, Pranith Kumar, Rohith Samineni, SR Rahul, Kaushal Sumedh, Dandu Geet Kamal Tej, and Vishal Rajapriya, "Machine learning based cloud integrated farming," In Proceedings of the 2017 International Conference on Machine Learning and Soft Computing, pages 1–6. ACM, 2017

[19] Min-Sheng Liao, Shih-Fang Chen, Cheng-Ying Chou, Hsun-Yi Chen, Shih-Hao Yeh, Yu-Chi Chang, and Joe-Air Jiang, "On precisely relating the growth of phalaenopsis leaves to greenhouse environmental factors by using an iot-based monitoring system," Computers and Electronics in Agriculture, 136, 125 (2017).

[20] Schubert Rodríguez, Tatiana Gualotuña, and Carlos Grilo, "A system for the monitoring and predicting of data in precision agriculture in a rose greenhouse based on wireless sensor networks," Procedia Computer Science, 121, 306 (2017).

[21] Ntihemuka Materne and Masahiro Inoue, "Iot monitoring system for early detection of agricultural pests and diseases,"

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:19:52 UTC from IEEE Xplore. Restrictions apply.

An Experimental Comparison of Different Object Tracking Algorithms

NS Raghava, Kushagra Gupta, Ishita Kedia and Anirudh Goyal

***Abstract*—This paper reviews and evaluates several state of the-art object tracking algorithms. It compares the various techniques on different parameters by implementing each one of them as it is important to check their performance and robustness. These algorithms are used to recognize, detect and track objects under the influence of different problems like occluded object, background clutter, illumination variations, and deformed objects or in shadowed regions. The tools used were python and its libraries including but not restricting to OpenCV, matplotlib, Numpy and others.**

***Index Terms*—Object tracking, OpenCV, python, compare;**

I. Introduction

OBJECT tracking in video streams and object detection and classification in still images, as well as videos, is an active and prominent research area in the field of computer vision. It has a plethora of applications in the real world which include surveillance, traffic control and analysis, face detection, medical image processing and many more [1]. While many advances and breakthroughs in this field over recent years have leaded to the development of a huge set of unique algorithms it still remains to be a very challenging problem.

Due to the enormous set of environmental factors and other variables, it is next to impossible to develop a universal tracking algorithm. Also, the selection of the most suited algorithm doesn't depend only on the concept of the algorithm but also on its implementation. Therefore, we compare some of the readily available algorithms on several challenging video sequences with different evaluation criteria using the OpenCV library [2].

The algorithms compared are easily found in the OpenCV library of python. It is a widely used library, with over 3000 algorithms embedded in it with all the needful structures and tools for object tracking algorithms.

N. S. Raghava is the Head of the Department of Electronics and Communication Engineering Department at Delhi Technological University, Delhi, India

Kushagra Gupta is with ECE department, Delhi Technological University, Delhi, India.

Ishita Kedia is with ECE department, Delhi Technological University, Delhi, India.

Anirudh Goyal is with ECE department, Delhi Technological University, Delhi, India.

This paper will help in observing the real-world implementations of the many algorithms found in this broadly used library. This evaluation and comparison discussed in this paper will not only be useful in selecting an appropriate method for specific applications but also be beneficial for researching and developing new tracking algorithms.

The remaining paper proceeds as follows: Section II gives a brief explanation of the different algorithms used. Followed by Section III which presents the various evaluation metrics against which the algorithms where compared. Section IV displays the results and the success and precision plots. Section V has the conclusion of the paper.

II. Algorithm Description

Describing the different algorithms implemented in this paper in brief to highlight the working mechanism along with a few important merits and demerits and Table I described the list of problems.

A. Tracking-Learning-Detection (TLD)

TLD is a framework used for tracking an arbitrary object in a video for a long time [3]. Since the three different subtasks of tracking, detection and learning are undertaken by different parts of the same process, it is described as a framework and not as a tracking method. The target object is simultaneously tracked and learned which will help make the algorithm robust against failures [4]. The tracker estimates the target's motion and predicts its position in the next frame according to the frame of reference. The Detector scans the image to restrict all appearances that have been observed before to a location. The Learning component estimates the detector's errors and produces numerous training examples to make sure that such errors are avoided in the future [5]. The framework aims to achieve improved tracking by preventing learning from disinformation. It achieves this by shutting down the learning task in case of occlusions or when the object is no longer in the frame. The detection component can detect the object again if it appears in the video afterward.

B. Boosting

Boosting is one of the older algorithms which is usually used for comparison to show an improvement in the latest ones. It's based on an online version of AdaBoost, which also powers the machine learning supporting Haar cascades.

978-1-7281-4988-2/20/$31.00 ©2020 IEEE

IEEE Advancing Technology for Humanity

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:26:18 UTC from IEEE Xplore. Restrictions apply.

This algorithm has to be trained rigorously with numerous positive examples (initial or starting bounding box provided by the user or the previous detection algorithms) and with an equal number of negative examples (patches around the actual object in the image). It works by working across every pixel of a given new frame but only surrounds a certain area of the previous location. Then the score is recorded of this classification algorithm for each comparison. The highest scoring part is labeled as the new location. As more new frames are added to this classifier it keeps getting updated.

C. Multiple Instance Learning (MIL)

With large datasets being used to solve complex issues today, research and interest in multiple instance learning has grown rapidly [6]. Another advantage is that it allows the use of data with weak labeling. As a result, it has been used for various applications such as drug activity prediction, text categorization and content based image retrieval and classification. It is a supervised learning algorithm used for cases in which the training data is not labeled completely. In supervised learning, all training instances are labeled individually. But, only a bag of instances is labeled in MIL [7]. In the case of binary classification, if at least one training instance in a bag is positive, the bag is positively labeled. Only if every training instance in a bag is negative, the bag is negatively labeled [8]. The objective of the algorithm is to classify instances or bags by using labeled bags as training data.

D. Kernelized Correlation Filter (KCF)

The KCF algorithm used the shift invariance property of Fourier transform [9]. This use not only simplifies the calculation of the correlation parameter but also makes it exceedingly fast. KCF is also able to bypass the issue of a huge amount of training data available from each frame of input due to its lightweight implementation. The large overlapping regions between various positive samples in Multiple Instance Learning are the property exploited by this tracker. KCF algorithm adopts and implements FFT and IFT which are fast and inverse Fourier Transform. By doing so the computation efficiency is increased of the algorithm [10].

The accuracy and speed of this algorithm can be easily be compared on accounts of performance to various different and latest algorithms in OpenCV library. Although it's reporting of tracking failure is an improvement over the previous versions it is unable to recover from full occlusions.

E. Median Flow

While using the median flow algorithm, an object is represented by a bounding box and then the object's movement is estimated between successive frames using sparse optical flow [11]. The algorithm assumes that the object is a combination of independent points which move in a synchronous manner and thus constitute the motion of the entire object. The algorithm is initialized by setting up the bounding box around the target object. A grid is then generated within the box. The points on the grid are then tracked between two consecutive frames. Next, the quality of the point predictions is estimated and each point is assigned an error. The best half of the predictions are used to estimate the displacement of the entire bounding box [12].

TABLE I
LIST OF PROBLEMS

Abbr.	Name	Description
BC	Background Clutters	Similarities in color or texture in the background near the target and the target itself.
DEF	Deformation	Non-rigid object deformation
FM	Fast Motion	The motion of the ground truth is larger than the limit
IPR	In-Plane Rotation	The target rotates in the image plane
IV	Illumination Variation	The illumination in the target region is significantly changes
LR	Low Resolution	The number of pixels inside the ground-truth bounding box is below limit
MB	Motion Blur	The target region is blurred due to the motion of the target or the camera
OCC	Occlusion	The target is partially or fully occluded
OPR	Out-of-Plane Rotation	The target rotates out of the image plane.
OV	Out-of-View	Some portion of the target leaves the view
SV	Scale Variation	The ratio of the bounding boxes of the first frame and the current frame is out of the range

III. EVALUATION METRICS

In this work, the success and precision rate of the algorithms are used for quantitative analysis and comparison. Additionally, the robustness of the algorithms has been evaluated using the TRE parameter.

TRE stands for Temporal Robustness Evaluation. TRE is a parameter to measure the tracker's robustness to initialization by varying the initialization by starting at different frames. Some tracker's performance may depend on the initialization, changing if the initialization is done at another start frame. The success parameter is used to define the intersections over union of the ground truth and the predicted bounding boxes while the precision metric defines the percentage of frame locations within a certain threshold distance from those of the ground truth [13].

A. Precision

This precision value is measured and tracked using a very common and well known metric called the Center Location error. Its definition is the average Euclidean distance between the manually labeled ground truths and the center locations of the tracked target [14]. Hence for precision evaluation, we utilized the scale of rectangle obtained from the tracker considering the ground truth using the following formula:

$$P = \frac{|r_t|}{|r_a|} \quad (1)$$

Here r_t is bounding rectangle obtained from the tracker and r_a is the bounding rectangle to represent the ground truth. The

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:26:18 UTC from IEEE Xplore. Restrictions apply.

overall performance for that sequence is measured using the average center location error over all the frames of one sequence. However, the output location could be arbitrary and the average center location error might not properly estimate the tracking performance if the tracker loses the target.

B. Success

Bounding box overlap is another evaluation metric. For evaluation of the success parameter we used the following formula for each frame is the video sequence:

$$S = \frac{r_t \cap r_a}{r_t \cup r_a} \quad (2)$$

Here r_t is the bounding rectangle obtained from the tracker and r_a is the bounding rectangle to represent the ground truth. The value thus obtained from the above formula is known as the overlap score which can be use to define success of the algorithm [2, 15-17]. The overlap score is essentially a ratio of the area of interception and area of union. Tracking algorithms which have a number of frames with an S value of greater than a given threshold are said to be successful. Although instead of basing it all on one threshold value we have also found out the values of Area Under Curve (AUC) for each algorithm for a better evaluation.

IV. Results

In this section we will be discussing all the overall results of the different algorithms employing tables and plots. All the results of the different problems were plotted for TRE for which all sequences were partitioned into 20 segments and therefore every tracker was tested on approximately 310,000 frames.

A. Overall Performance

The overall performance can be a combined output of the graphs measuring precision and success. For success plots, AUC scores were employed to evaluate the trackers. Both the precision and success plots have a few different results. This is because both of them are based upon different metrics. Mainly success results were used to rank but we used the precision plots as an auxiliary source.

Based on the criterion of success we got unexpected results as TLD algorithm, considered to be amongst the best, and was not able to perform well. Mainly two reasons can be the cause of such a result one being the mere complexity of the algorithm and the other is the insufficient implementation in OpenCV, Also it can be seen from the time comparison defined by the parameter of Frames Per Second (FPS) that the MIL is the slowest algorithm while Median Flow turns out to be the fastest. The comparison is shown in Table IV. Once again a surprising result was obtained due to the TLD's slow speed which is believed to be caused by its implantation using OpenCV which has a lack of optimizations.

To gain a better overview a summarized form of the success scores and the precision scores for each algorithm across the different attributes is given in Table II and Table III.

TABLE II
PROBLEMS' SUCCESS TABLE

Abbr.	MIL	KCF	TLD	MF	Boosting
IV	0.332	0.332	0.265	0.252	0.381
SV	0.395	0.435	0.223	0.305	0.404
OCC	0.318	0.361	0.267	0.266	0.318
DEF	0.420	0.489	0.245	0.313	0.416
MB	0.405	0.480	0.227	0.317	0.382
FM	0.384	0.466	0.256	0.351	0.389
IPR	0.283	0.263	0.292	0.237	0.333
OPR	0.388	0.464	0.275	0.311	0.391
OV	0.405	0.470	0.262	0.358	0.393
BC	0.334	0.374	0.263	0.236	0.315
LR	0.375	0.419	0.275	0.323	0.378

TABLE III
PROBLEMS' PRECISION TABLE

Abbr.	MIL	KCF	TLD	MF	Boosting
IV	0.4	0.385	0.356	0.275	0.381
SV	0.509	0.591	0.334	0.374	0.506
OCC	0.410	0.439	0.383	0.311	0.365
DEF	0.557	0.677	0.345	0.381	0.523
MB	0.532	0.665	0.320	0.364	0.472
FM	0.519	0.667	0.383	0.429	0.510
IPR	0.384	0.329	0.470	0.260	0.463
OPR	0.522	0.655	0.386	0.381	0.506
OV	0.560	0.680	0.381	0.446	0.522
BC	0.315	0.396	0.283	0.221	0.319
LR	0.546	0.646	0.406	0.412	0.582

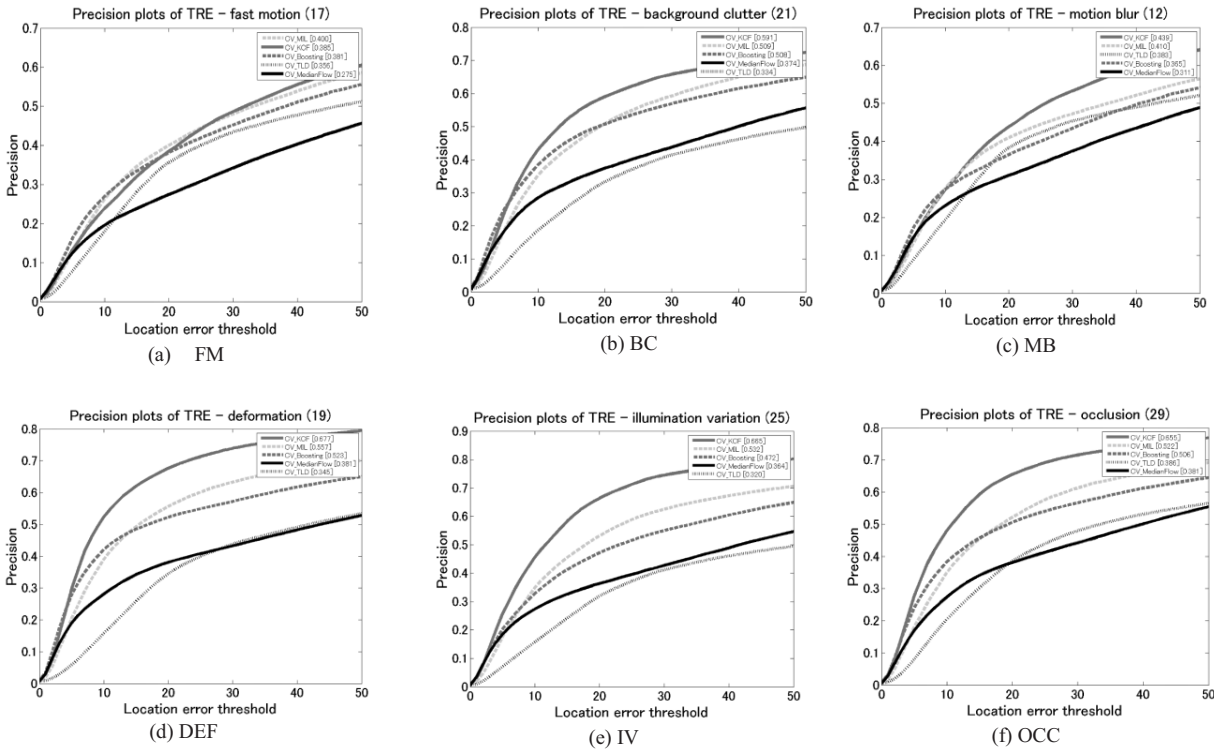
TABLE IV
RUN-TIME COMPARISON

Abbr.	MIL	KCF	TLD	MF	Boosting
FPS	16.35	230.26	18.77	300.15	30.20

B. Attribute Based Performance

By analyzing the attributes of all video sequences, we created different subsets. Each of the subsets had different dominant attributes which can be considered as problems which have to be overcome during object tracking. Therefore, the examination of the performance of each tracker for all individual challenging problems became easy.

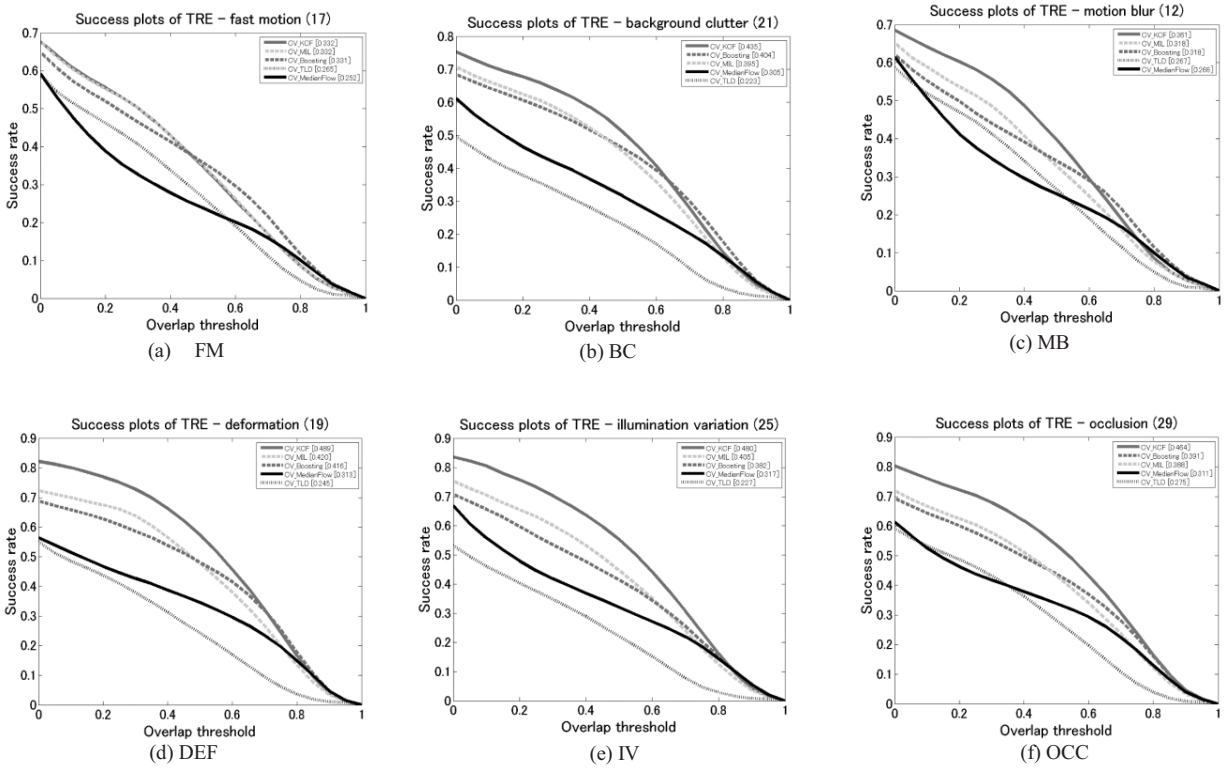
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:26:18 UTC from IEEE Xplore. Restrictions apply.



(a) FM (b) BC (c) MB

(d) DEF (e) IV (f) OCC

Fig. 1. Precision Plots of all trackers in different problems



(a) FM (b) BC (c) MB

(d) DEF (e) IV (f) OCC

Fig. 2. Success Plots of all trackers in different problem

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:26:18 UTC from IEEE Xplore. Restrictions apply.

In the precision plots shown in Fig. 1, we observe KCF to be the most successful algorithm in overcoming most of the problem parameters upon which the experiments were performed. Except for the case of Fast motion where MIL algorithm turned out to be more effective than KCF.

Also, TLD outperforms KCF in the case of low resolution. For the plots indicating the success Vs the overlap threshold as shown in Fig. 2, the legend carries the AUC score which summarizes and ranks the trackers being compared. The success plots also observed similar results to those of the precision plots. KCF turned out to be the best in overcoming all attributes but Fast motion and low resolution.

V. Conclusion

When all the algorithms were tested under the same environment, it can be concluded that BOOSTING Tracker is quite slow and doesn't work very well. It should be noted that it has only been included for legacy reasons. The MIL Tracker is found to be wanting against other algorithms especially in reporting failure. It was interesting to observe the TLD tracker as it was incredibly prone to false-positives, which was not initially expected, thus not making a good choice when using the OpenCV object trackers. Median-Flow Tracker was the fastest tracker to be tested in our environment which did a considerably good job when it came to reporting failures. However, if the object changes their appearance quickly enough, or there is a too large jump in object motion, the object changes their appearance quickly enough, or there is a too large jump in object motion, the tracker fails to perform. The tracker should not be used for fast-moving objects. The Kernelized Correlation Filters (KCF) tracker was one of the quicker trackers, faster than BOOSTING, MIL, and TLD tracker. It was noted that similar to MIL Tracker it was prone to occlusion but was much more reliable in its performance, making it one of the most favorable trackers that can be used. The future scope of this research includes a comprehensive examination of the application of these algorithms, along with the calculation of their efficiencies and the use of different techniques.

Acknowledgment

This research paper is supported by the Electronics and Communication Engineering Department, Delhi Technological University, New Delhi.

References

[1] Q. Wang, F. Chen, W. Xu, and M.-H. Yang, “An experimental comparison of online object-tracking algorithms,” *Wavelets and Sparsity XIV*, Aug. 2011.

[2] P.Janku, K.Koplik, T.Dulik, and I.Szabo, “Comparison of tracking algorithms implemented in OpenCV,” *MATEC Web of Conferences,* vol. 76, p. 04031, 2016.

[3] Z. Kalal, K. Mikolajczyk and J. Matas, "Tracking-Learning-Detection," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 7, pp. 1409-1422, July 2012.

[4] J. Hu, S. Hu and Z. Sun, "A real time dual-camera surveillance system based on tracking-learning-detection algorithm," *2013 25th Chinese Control and Decision Conference (CCDC)*, Guiyang, 2013, pp. 886-891

[5] *Jig et al*., "A Tracking-Learning-Detection (TLD) method with local binary pattern improved," *2015 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Zhuhai, 2015, pp. 1625-1630.

[6] M.-A. Carbonneau, V. Cheplygina, E. Granger, and G. Gagnon, “Multiple instance learning: A survey of problem characteristics and applications,” *Pattern Recognition*, vol. 77, pp. 329–353, 2018.

[7] O. Maron, and T. Lozano-Pérez. "A framework for multiple-instance learning." In *Advances in neural information processing systems*, 1998, pp. 570-576.

[8] B. Babenko, M. Yang and S. Belongie, "Visual tracking with online Multiple Instance Learning," *2009 IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL, 2009, pp. 983-990.

[9] V.lehtola, H.Huttunen, F.Christophe, and T.Mikkonen “Evaluation of Visual Tracking Algorithms for Embedded Devices,” *Image Analysis Lecture Notes in Computer Science,* pp. 88-97, 2017.

[10] C. Zhang *et al*., "Output Constraint Transfer for Kernelized Correlation Filter in Tracking," in *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 4, pp. 693-703, April 2017.

[11] Varfolomieiev and O. Lysenko, “An improved algorithm of median flow for visual object tracking and its implementation on ARM platform,” *Journal of Real-Time Image Processing*, vol. 11, no. 3, pp. 527–534, 2013.

[12] Z. Kalal, K. Mikolajczyk and J. Matas, "Forward-Backward Error: Automatic Detection of Tracking Failures," *2010 20th International Conference on Pattern Recognition*, Istanbul, 2010, pp. 2756-2759.

[13] X. Dong, J. Shen, W. Wang, Y. Liu, L. Shao and F. Porikli, "Hyperparameter Optimization for Tracking with Continuous Deep Q-Learning," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, 2018, pp. 518-527.

[14] Y. Wu, J. Lim and M. Yang, "Online Object Tracking: A Benchmark," *2013 IEEE Conference on Computer Vision and Pattern Recognition*, Portland, OR, 2013, pp. 2411-2418.

[15] S. Moon, J. Lee, D. Nam, H. Kim and W. Kim, "A comparative study on multi-object tracking methods for sports events," *2017 19th International Conference on Advanced Communication Technology (ICACT)*, Bongpyeong, 2017, pp. 883-885.

[16] Bibi and B. Ghanem, "Multi-template Scale-Adaptive Kernelized Correlation Filters," *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, Santiago, 2015, pp. 613-620.

[17] W.Chen, K.Zhang, and Q.Liu, “Robust visual tracking via patch based kernel correlation filters with adaptive multiple feature ensemble,” *Neurocomputing* vol.214, pp.607-617, 2016.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:26:18 UTC from IEEE Xplore. Restrictions apply.

Classification of EEG signal using LSTMs under Audiovisual Stimuli

Tejas Agarwal, Shubham Raturi, TK Vybhav and Mukhtiar Singh

***Abstract*—This paper presents the latest developments in the field of Electroencephalogram (EEG) signals classification and further classifies the BCI data generated due to Audio/Visual Stimuli using a variety of Long Shot Term Memory's (LSTM) deep learning models. The techniques involved in the classification are Vanilla LSTM, Bidirectional LSTM and Stacked LSTM. An arduous task at hand was to preprocess the irregular data using Independent Component Analysis (ICA) and using MATLAB's EEGLAB software. The dataset was recorded from readings generated from users being subjected to tones and flashes at regular and irregular intervals. Another analysis was performed to find out the best optimizer for BCI data. It was seen that the most efficient classifier was Bidirectional LSTM pointing to 84% accuracy.**

***Index Terms*—Brain Computer Interfaces, Deep Learning, Electroencephalograph (EEG) signals, Long short-term memory (LSTM), Machine Learning, Signal Processing**

I. Introduction

ELECTROENCEPHALOGRAM (EEG) based Brain Computer Interfaces (BCI) have gained a massive interest as a direct, external communication pathway between an external device and the human brain. The procedure is applicable to almost every person with a healthy brain and is noninvasive, safe and easy to apply.

Human beings can now control machines or devices without any muscle movement involved, by using a BCI as an alternative communication system between the two. The brain generates electroencephalography (EEG) signals [1], which are converted into control commands by BCI using machine learning techniques. People suffering from neuromuscular disease find great use of BCI as numerous functions, like commanding a computer, using common communication devices, can now be performed by using their brain thought [2].

A computer is now capable of controlling a wide range of electronic equipment including wheelchairs, gateways, lamps and home appliances. In the earlier stages, BCI was used in enhancing different life aspects for the severely disabled by providing them assistive technology [3]. But today it finds substantial use in non-medical applications such as controlling a personal computer or smart home appliances, whose users can be both disabled and otherwise.

Tejas Agarwal, Shubham Raturi, TK Vybhav, Mukhtiar Singh, Department of Electrical Engineering , Delhi Technological University , New Delhi, India

Every BCI consists of five stages which it goes through. The first is measurement of brain activity, second is preprocessing the data, third is extraction of features, fourth is classification and final is control interface by which meaningful information for output devices is extracted from the classified information. [4] Classification techniques were used to classify the features which were obtained from the feature extraction stage and effective classification results which agree with the user's intentions is only achievable through a choice of good distinguishing features at this stage. Due to different experiment settings and different environments, the results of different techniques cannot be compared effectively [5]. Also, the main processes which favorably affect the performance accuracy are preprocessing, feature extraction and classification stage. So, to get a high accuracy result, the appropriate technique in these stages is a crucial parameter to determine [6].

This paper lists down the conventional techniques being used for classification of EEG signals and the processes involved. Further, they also compare three deep learning techniques of LSTM and generate metrics to determine the efficiency of each in this problem of classification. The objective in this paper is to find the most efficient classifier for motor imagery BCI. The classifiers are evaluated using BCI dataset from Dr. Ramesh Balasubramaniam's laboratory.

The remainder of the paper is organized as follows: Section II defines the classifications methods. Data processing and results are defined in section III &IV. The paper is concluded in section V.

II. Classification Methods

A. SVM

Vladimir N. Vapnik invented the Support Vector Machine (SVM) algorithm in 1963; however, it was revised in 1992 by Boser, Guyon and Vapnik to attain nonlinear classification by using the kernel trick. SVM uses a hyperplane to distinguish classes; the hyperplane is selected which supports margin maximization the most, while using regularization parameters to correct the algorithm for outliers and errors while training the dataset. It can create linear decision boundaries or use kernels to achieve non-linear decision boundaries. [7] This kernel technique helps in analyzing BCI application data as it allows the data to be mapped to a higher dimensionality. Overfitting is also avoided by the use of regularization terms and the concept of margin maximization.

978-1-7281-4988-2/20/$31.00 ©2020 IEEE

IEEE Advancing Technology for Humanity

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:25:05 UTC from IEEE Xplore. Restrictions apply.

B. K- Nearest Neighbor

K Nearest Neighbor is a nonparametric algorithm used for applications requiring regression and classification problems. A feature vector is assigned to class with the help of the nearest neighbor, using the most effective class from the training set. In BCI the metric distance is used for calculating the nearest neighbors, due to presence of high k value and large training set kNN is able to produce non-linear decision boundaries which are extremely sensitive to the curse of dimensionality which makes them not suitable for most of the BCI applications other than the ones with low dimensional feature vectors. The k value can be calculated by using weighted distances and cross validation.

C. Convolution Neural Networks

Progress in the field of Convolutional Neural Networks (CNNs) has followed a process similar to vision processing of organisms. It began around the 1950s but the first convolutional network was introduced in 1987 by Waibel et al achieving shift invariance in Time Delay Neural Networks (TDNN)[8]. CNNs are deep neural networks involved in solving complex imagery problems by detecting abstract features. They are biologically inspired Multi-layer perceptrons (MLP) and employ minimum preprocessing compared to other image classifiers. More importantly, they learn task related filters without any pre-fed information and offer the perks of a feature extractor and classifier in a single algorithm, In their paper, Z. Tang et al [9] discusses EEG Motor Imagery(MI) Classification using a 5 layer model based on spatio-temporal characteristics and presents an effective and non-evasive way in BCI applications.

D. Long Short Term Memory

LSTM or Long Short Term Memory is a kind of RNN and was first proposed by Hochreiter and Schmidhuber in 1997. LSTMs have widespread uses in real life problem because of their property to selectively remember patterns for long durations of time especially in Sequence Prediction.

The idea behind their development was to tackle the exploding and vanishing gradient problems. Now, its feedback connections enable the user for classifying and making predictions on data sequences such as EEG time series.

There are a variety of LSTM networks like the Gated Recurrent Unit (GRU), peephole connection LSTM, developed from different perspectives. Wang et al use 1d-AX to extract effective signal representation from EEG data and apply channel weighting technique to improve the effectiveness in EEG classification of motor imagery tasks.

III. Data and Preprocessing

A. Dataset

Dr. Ramesh Balasubramaniam conducted a two part experiment the premise of which was to expose the subject to different stimuli. The two different stimuli used were audio and visual in nature, to be more precise they were tones and flashes. Part 1 of the experiment had the subject experience a block of isochronous, uninterrupted tones and flashes that total 140 per stimuli; for part 2 the subject experienced another block that had occasional omissions in either the tone or flash, but still total up to 140 per stimuli. The frequency of the stimuli was varied and the subjects were asked if they felt any change in the speed of the same.

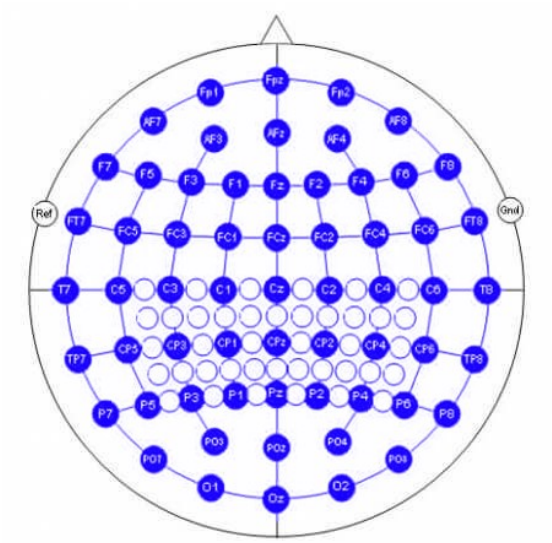


Fig. 1. 32 Channel EEG Cap

Brain waves are measured using electroencephalography (EEG) technique, the amount of data gathered depends upon the number of channels measured by the waveguard (EEG cap). Dr. Ramesh Balasubramaniam's experiment made use of a 32 channel cap; its basic specification and design with respect to the different regions of the brain are given as follows.

The two main regions of the brain i.e. somatosensory cortex and motor cortex are mapped using the 32 channel EEG cap shown in Fig. 1. Odd numbers to the left and even on the right with the letters F,T,C,P,O,Z representing frontal lobe, temporal lobe, central lobe, parietal lobe, occipital lobe and electrode respectively.

The recorded EEG signal shows that the subjects' brain waves change in response to the stimuli. This data is stored as a .set file with its corresponding .fdt file. Using EEGLab toolbox for MATLAB these .set files can be opened and viewed. The recorded data shows us the EEG data with corresponding event codes and event latencies.

B. Preprocessing

Electric signals are recorded from the surface of a subject's scalp, which forms the raw data for EEG. This collected raw data further contains components of noise from interference due to muscular or eye movement, and underlying electronic signals within the brain of the subject. To process this EEG data, a common approach is to resolve the 32 channels measured externally into independent signals internally. This is done popularly by a technique called Independent Component Analysis (ICA). [10]

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:25:05 UTC from IEEE Xplore. Restrictions apply.

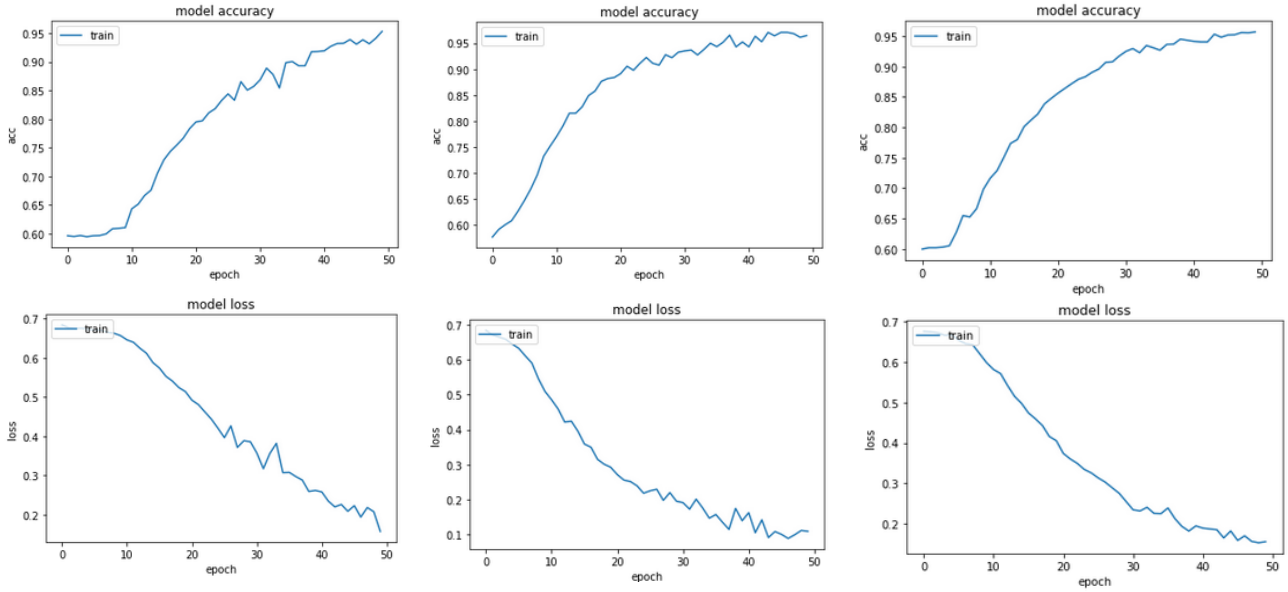


Fig. 2: Row 1: Model Training Accuracy Graphs; Row 2: Model Training Loss Graphs,
From L to R: Vanilla LSTM, Bidirectional LSTM, Stacked LSTM

This technique successfully represents high dimensional data into a smaller set of components and has commonly been developed in numerical linear algebra. Another technique called Principal Component Analysis is also used for this purpose but it either produces orthogonal components or requires the underlying signals to be Gaussian. For datasets like EEG, where data contains multiple simultaneous recordings, ICA is preferred over PCA. EEG data will be translated into lower dimensions by ICA using the MATLAB codebase EEGLab [11], in the first step.

A recurrent neural network (RNN) will be developed after the EEG data is decomposed into independent components, which will classify the subject data based on whether a visual or auditory stimulus was received by the individual. A small dataset (4 subjects in each class) will be provided to the RNN to train it and the remaining 24 subject readings (12 in each class) will be classified into their appropriate class (auditory or visual).[12] Fully connected networks employing rectified linear (ReLu) units will be explored, keeping in mind that there exist many choices for neural networks.

Low dimensional signals must be identified and exploited to promote real time classification and analysis of subject behavior from EEG data. Characteristic signals in the independent components were attempted to identify, indicative of either stimulus, after completing the classification of audio or visual stimuli of the individual.[13] Common components or inter component relationships indicating subject exposure were obtained by using cluster analysis.

C. *Experimental Setup*

The two MATLAB datasets were loaded into two CSV files to visualize the data as just numeric values instead of graphs and charts. This allowed us to see exact values, at exact time stamps of the events, which ensured that the model was being trained on the correct events and times. The CSV file containing the EEG signals are 1.7GB in size, and with the time scale being on Milliseconds, this amount of data was sufficient in training the model.

D. *Model Description*

The dataset was tested for classification on three different models of LSTM namely, Vanilla LSTM, Bidirectional LSTM and Stacked LSTM.

General model included 3 layers shown in Fig. 3, they are:

- Input layer
- LSTM or wrapper (Bi)
- Dropout
- Dense (Sigmoid)

The ‘sequential ’model had 100 neurons and ran for 50 epochs each. Loss function used was ‘Binary cross entropy’. The Dropout layer randomly drops neurons at rate 0.5 of the input units, which helps overfitting.

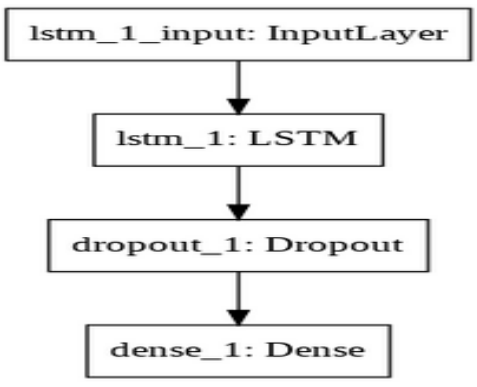


Fig. 3. Model Layer Layout

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:25:05 UTC from IEEE Xplore. Restrictions apply.

IV. Results

The results from classifying the Audio/Visual data with three different models are presented in Fig. 2 above. The first row depicts the accuracies of the model's Vanilla LSTM, Bidirectional LSTM and Stacked LSTM respectively. The second row follows with the Model loss for the three methods. The program was run on these LSTMs multiple times to get a minimum and maximum limit or range for the model. The accuracy mentioned below is of a test dataset that is 20% of the total number of samples.

For Vanilla LSTM, the highest accuracy achieved was (80.38%) while a range of 75% – 80% was seen on consecutive runs of the model.

TABLE I
COMPARISON OF PERFORMANCE OF LSTMS

Model	**Accuracy**	**Time/epoch**
Vanilla LSTM	80.38%	180s 73ms
Bidirectional LSTM	84.31%	240s 119ms
Stacked LSTM	82.56%	252s 115ms

Bidirectional LSTM saw the highest accuracy among all three of (84.31%) and maintained a range of 78% - 84% at all times. Stacked LSTM performed excellently while training and gave an accuracy of (82.56%), with a range of 76% - 82%.

The model was also run by modifying the optimizer to 'Adam' and 'sgd' giving accuracy around 61.92% and 63.12% respectively. Overall, they performed poorly with random peaks and valleys in the graph.

V. Conclusion

Thus, we successfully classified BCI data using a variety of LSTM techniques. Our experiment proves that RNNs are novel methods of time series classification and also provide high accuracies as in the case of Bidirectional LSTM (84%). Further advancements would include processing of real-time EEG data and employing it in the medical as well as the research domain. Comprehending brain data is like opening doors to telekinesis and to a new era of human cognition.

References

[1] Electroencephalography (EEG) Scan Definition & Test Results. (n.d.). Retrieved November 19, 2017, from https://www.emedicinehealth.com/electroencephalography_eeg/article_em.htm#electroencephalography_eeg_introduction.

[2] Géron, A. (2017). Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems.

[3] Liu jingwei, Cheng Yin, Zhang Weidong. "Deep Learning EEG Response Representation for Brain Computer Interface" in 34th Chinese Control Conference

[4] July 28-30, 2015, Hangzhou, China.B. Graimann, B. Allison, and G. Pfurtscheller, "Brain–Computer Interfaces: A Gentle Introduction," in Brain-Computer Interfaces, B. Graimann, G. Pfurtscheller, and B. Allison, Eds., ed: Springer Berlin Heidelberg, 2010, pp. 1-27.

[5] Liu jingwei, Cheng Yin, Zhang Weidong. "Deep Learning EEG Response Representation for Brain Computer Interface" in 34th Chinese Control Conference

[6] P. Sajda, A. Gerson, K. R. Muller, B. Blankertz, and L. Parra, "A data analysis competition to evaluate machine learning algorithms for use in brain-computer interfaces," *IEEE Transactions on Neural Systems and Rehabilitation*

[7] Gary N.Garcia, Touradj Ebrahimi and Jean-Marc Vesin, "Support Vector Eeg Classification In The Fourier and Time-Frequency Correlation Domains", IEEE EMBS

[8] Alexander Waibel, "Phoneome Recognition using Time Delay Neural Networks". IEEE Transactions on Acoustic Speech and Signal Processing, Vol. 37. No. 3, March 1989

[9] Zhichuan Tang, Chao Li , Shouqian Sun, "Single-trial EEG classification of motor imagery using deep convolutional neural networks " in Optik 130 (2017) 11–18

[10] C. Corradino and M. Bucolo, "Automatic preprocessing of EEG signals in long time scale", 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)

[11] Delorme & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of neuroscience methods, 134(1), 9-21.

[12] M. Z. Ilyas, P. Saad and M. I. Ahmad, A. R. I. Ghani "Classification of EEG Signals for Brain-Computer Interface Applications: Performance Comparison"

[13] Sheng Li and Hanxin Feng, "EEG Signal Classification Method Based on Feature Priority Analysis and CNN", 2019 International Conference on Communications, Information System and Computer Engineering (CISCE)

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:25:05 UTC from IEEE Xplore. Restrictions apply.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/343636160

COVID-19 Outbreak Prediction using Quantum Neural Networks

Conference Paper · August 2020

CITATIONS
0

READS
227

2 authors:



Pranav Kairon
Delhi Technological University
2 PUBLICATIONS **0** CITATIONS
SEE PROFILE



Siddhartha Bhattacharyya
Christ University, Bangalore
325 PUBLICATIONS **1,154** CITATIONS
SEE PROFILE

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

Project Feature Selection for large Scale Data sets (Big Data) View project

Project Call For Book: Swarm Intelligence for Cloud Computing View project

All content following this page was uploaded by Pranav Kairon on 13 August 2020.

The user has requested enhancement of the downloaded file.

COVID-19 Outbreak Prediction using Quantum Neural Networks

Pranav Kairon[1] and Siddhartha Bhattacharyya[2]

[1] Delhi Technological University, Bawana, Delhi, India,
pranavkairon_bt2k17@dtu.ac.in
[2] CHRIST (Deemed to be University), Bangalore, India,
dr.siddhartha.bhattacharyya@gmail.com

Abstract. Artificial intelligence has become an important tool in fight against COVID-19. Machine learning models for COVID-19 global pandemic predictions have shown a higher accuracy than the previously used statistical models used by epidemiologists. With the advent of quantum machine learning, we present a comparative analysis of continuous variable quantum neural networks (Variational circuits) and quantum backpropagation multilayer perceptron (QBMLP). We analyze the convoluted and sporadic data of two affected countries, and hope that our study will help in effective modeling of outbreak while throwing a light on bright future of quantum machine learning.

Keywords: COVID-19, Corona virus, Quantum machine learning, Quantum neural network

1 Introduction

The Novel Corona virus disease/COVID-19 was first detected in Wuhan, China on 31st December 2019. After the 1918 H1N1 influenza, COVID-19 has been accounted for as the most pernicious respiratory infection to affect the population worldwide [1]. With meteoric spread it has taken lives of half a million people and around 13 million confirmed cases bringing lives of millions to a standstill. According to the World Health Organization report 170 countries now have adumbrated at least one case as on July, 2020. While, 4.5% is the mortality rate of this deadly disease, for the age group 70-79 this has ascended to be 8.0% while for those above 80 it is soaring 14.8%. People already suffering from heart disease, diabetes are especially at higher risk, the fact that they are older than 50 only exacerbates the situation [2]. Middle East Respiratory Syndrome (MERS), Severe Acute Respiratory Syndrome (SARS) and COVID-19 all belong to a family of severe acute respiratory syndrome, called as coronavirus. The symptoms are self evident within 2-14 days [3] and have been signified by gamut of fever, cough, shortness of breath to pneumonia. Flattening the epidemic curve is key part in managing this pandemic since there is no commercially available vaccine against COVID-19. Although there is some hope as researchers in Russia and Oxford University have completed the human trials of the vaccine, which show

positive results.
Given the effectual predictions in healthcare, machine learning has proven itself fundamentally. After intensifying patterns and expertise of radiologists, machine learning is now helping them to predict the disease and diagnose it earlier. Also with availability of meager data on COVID-19 the role of data scientists has increased drastically Since they have to integrate the data and then release the results which will help in taking unerring decisions. [4].
Quantum computing [5] is an amalgamation of quantum physics and computer science. Originally proposed in the 1970s, the research remained mostly theoretical until early 2000s with advent of Shor's and Grover's Algorithms [6]. We are presently in the NISQ (Noisy intermediate Quantum) era of quantum computing which implies that the presently available quantum processors are small and noisy. Since many countries are investing heavily on quantum technologies, the growth rate has been accelerated in past two years. Quantum computers supersede classical computers by harnessing quantum effects for computational vantage providing polynomial and even exponential speedups for specific problems. With Moore's law coming to an end, quantum computers become even more eminent since they are not made up of small transistors but rather quits. On the other hand we are living in an age of data and machine learning provides robust models for predictions, classification and organisation tasks.

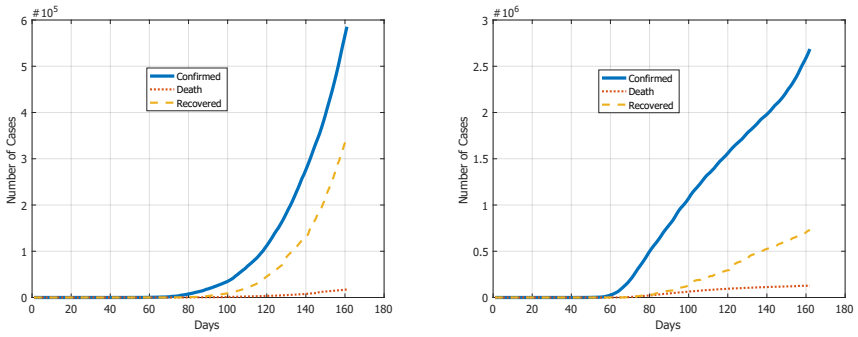
2 Motivation

Presently available artificial neural network models provide robust models for plethora of problems ranging from stock prediction and statistical modeling to image recognition, data mining [7]. However, quantum neural networks are still somewhat unexplored in terms of their applicability to eclectic and pragmatic problems. In terms of convergence rate and fit ability, Quantum Backpropagation Neural Networks have outperformed classical ANNs, since they can harness the mathematical advantage of complex numbers [8] using Quantum phase door and two negates controlled controlled Quantum neuron model. Continuous variable Quantum neural networks [9] are similar to variational circuits in that they are free parameter dependent quantum algorithms. They are gaining attention as an alternate way to characterize NISQ era devices. Given the guessing abilities of quantum neural systems, these networks are agile to cost degradation.These properties have roused us to apply these network models to regressional tasks like predicting rise in cases for COVID 19 and compare them against the intelligent paradigmatic classical ANNs.

3 Proposed Methodology

The experimental data has been collected from [15] for India and USA. We have considered a live time-series dataset from 30th January till 1st July 2020. It provides the number of confirmed cases, number of deaths and number of people who have recovered from the pervasive infection. We can find additional two features

i.e. the number of active cases which is given by Active cases=No.of Confirmed Cases - (No.of deaths + No. of recovered patients) and New cases added each day which is just the difference of active cases on the nth day and $(n-1)$ th day. The global data for all the three columns is plotted in Figure 1. The pervasive nature of the infection is evident from the exponential curves in Figure 1. In this work, Quantum Neural Networks (QNNs) have been employed to predict rise in cases. QNNs are seldom applied to regressional tasks [10]. We consider two kinds of quantum neural networks. First being, a Quantum Backpropagation Multilayer Perceptron (QBMLP) which utilizes the superposition feature and complex numbers [8][11]. Secondly, we employ a Continuous Variable Quantum Neural Network (CVQNN) provided by Xanadu in their Pennylane and Strawberry fields package [9]. CVQNNs are variational circuits that use Gaussian and non-Gaussian gates to perform affine and non linear transformations. Both the QBMLP and CVQNN can be generalized to any number of hidden layers, and are characterized by the Tanh and Kerr activation functions, respectively. The accuracy for both the models in finding out the rise in number of cases and deaths due to COVID-19 in India and USA is tested. Moreover, a comparative study has been made between both the networks using a two tailed statistical t-test. CVQNN outperforms QBMLP in some cases whereas an opposite effect is seen in some, as can be seen from the two tailed t-test.



(a) India COVID-19 cases (b) USA COVID-19 cases

Fig. 1. Total number of cases for all the countries from January 30th to 1st July 2020

4 Models

Deep feed forward neural networks have shown tremendous results on problems like image processing, stock predictions etc. They owe their success to affine transformations done through hidden layers, which are the building blocks of a mulitlayer perceptron, and activation functions that introduce non linearity to

map input x to output y. A typical architecture can be represented by

$$L : x \rightarrow y = \phi(Wx + b) \tag{1}$$

for l input variables and m output variables, W is a $l \times m$ matrix and b is constant vector of length l, where $W, b \, \epsilon R$. ϕ is usually chosen from a pool of functions, it can be sigmoid, tanh, $ReLU(x) = max(0, x)$. Inputs for all the layers form the output for the next layers, and the weights are optimized for the defined task on basis of hyper parameters adjusted by the user, which include the number of layer wise input variables (features) and number of layers (*depth* d).

4.1 Fuzzified Quantum Backpropagation Multilayer Perceptron

Consider a dataset of N features that can be represented as $x = (x_{1,}x_{2,}....x_N)$. The input data can be squashed into the range $[0, 1]$ using fuzzy linear membership function [11] with 1 signifying the highest value whereas 0 lies on opposite end of the spectrum. Membership value is allotted to all the data points in between using linear scaling rule [8].
Given the inputs x, we define a function $f(x)$ as

$$|x\rangle = cos(x)\,|0\rangle + sin(x)\,|1\rangle \equiv cos(x) + i.sin(x) \tag{2}$$

The weights which are modeled as rotation gates, $R(\theta_i)$ to the input qubits, $|x_i\rangle$ with input bias λ. Hence, the modified neural network equation takes the form

$$u = \sum_{i=1}^{n} R(\theta_i)\,|x\rangle - |\lambda\rangle \tag{3}$$

Now, instead of applying just a non linear function as in the classical case, the transformed output is given by

$$y = \frac{\pi}{2} g(\delta) - \arctan(\frac{amplitude(|1\rangle)}{amplitude(|0\rangle)}) \tag{4}$$

where, $g(\delta)=\frac{1-e^{-2\delta}}{1+e^{-2\delta}}$ is the tanh function.
Similar to the classical Multilayer Perceptron, the QBMLP is designed with the first layer accepting inputs in form of qubits, followed by any number of hidden layers, ending at an output layer with a single neuron (for regressional tasks). The qubits for QBMLP are designed by transforming the squashed input x into the range of $[0, \frac{\pi}{2}]$ so that they become suitable angles for rotation gates, which after can be transformed to

$$\left|\frac{\pi}{2} x_i\right\rangle = cos(\frac{\pi}{2}x)\,|0\rangle + sin(\frac{\pi}{2}x)\,|1\rangle \tag{5}$$

The output of the model is a quantum state expressed as a complex number in Euler's form. To extract the physical meaning of this output, either a quantum measurement can be done and the probability of getting $|0\rangle, |1\rangle$ is considered as

the output or equivalently the final output can be obtained as square of either the real or the imaginary part of the output. In this paper, we have taken complex part of output which is the same as taking probability of getting $|1\rangle$ as the final result. The necessary changes need to be made in the analysis of backpropagaton if one decides otherwise.

$$\mathcal{O} = Probability(|1\rangle) =\mid Im(y) \mid^2 = sin^2(y) \tag{6}$$

Interested readers may refer to [11] for details on the architecture and operation of the network architecture.

4.2 Continuos Variable Quantum Neural Network

Pennylane is a quantum macine learning python library provided by Xanadu[12], working on the principle of quantum differentiable programming. It provides an efficient platform by combining quantum simulators and classical machine learning, that helps users in traning various circuits. [13]. The Strawberry fields in a photonic quantum computing library used to solve plethora of problems such as boson sampling, graph optimiation etc. This has been used to construct a photonic neural network model having continuos variable gates. Variational circuits [14] behave similar to neural networks in that, there is a definite input (input quantum state) which is embedded into the circuit using a suitable embedding, weights (circuit parameters) that need to be learnt by the model and quantum measurement of an observable say A, which generates a classical output on which training rules are applied for parametric learning. A three qumode CVQNN architecture is shown in Figure 2 (a). They are quite useful algorithms for NISQ devices in prblems such as quantum chemistry, optimization etc.
This continuous variable quantum neural network actually contains classical neural networks within itself. The use of continuous variable Gaussian and non Gaussian operations such as displacement operators, squeezing gates, Kerr gates etc, attribute novelty to this approach given the fact that it doesn't contain quantum features such as entanglement and superposition in contrast with other variational quantum circuits. If the input dataset has N features (columns), it can be encoded into the circuit using displacement operators. Let the classical input x is given by $x = (x_1, x_2...x_N)$ and the initial quantum state of circuit is given by $|x\rangle = \bigotimes_{i=1}^{N} |x_i = 0\rangle$. Then the displacement operators, $D(x)$ can be applied to all qumodes, to encode classical input x such that

$$|x\rangle = \bigotimes_{i=1}^{N} D(x_i)|x_i = 0\rangle \tag{7}$$

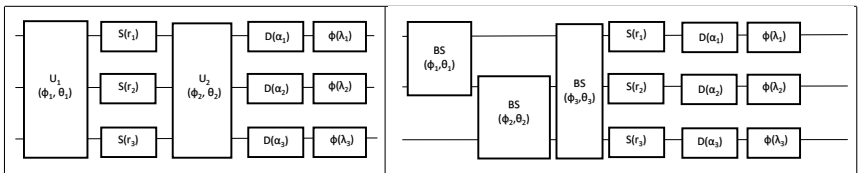
Just like in classical feed forward neural networks, any layer takes inputs from the output of the preceding layer, results are extracted by applying homodyne measurements on each qumode. A similar type of transformation has been used, by using continuous variable quantum analogs. Information about the various CV analogs applied in the quantum neural network can be found in Table 1.

Combining all of the above, in order to obtain the our full affine transformation $D \circ U \circ S \circ U \left|x\right\rangle = \left|Wx + d\right\rangle$, we can simulate a classical neural network layer which if broken down is basically a transformation given by equation 8, and the post proccessing of the output is similar to what is done in all neural network models, i.e. paramters are updated on basis of training rules as

$$L(z) = \phi(Mz + \alpha) \tag{8}$$

Table 1. Various CV quantum gates used to simulate affine transformation

Classical	CV analog	Transformation
Weight Matrix : W	Interferometer :$\hat{\mathcal{U}}$	$\hat{\mathcal{U}}\|x\rangle = \bigotimes_{i=1}^{N} \|\sum_{j=1}^{N} C_{ij}x_j\rangle = \|C\mathbf{x}\rangle$
	Squeezing gate : $\hat{\mathcal{S}}$	$\hat{\mathcal{S}}(r)\left\|x\right\rangle = e^{-\frac{1}{2}\sum_i {r^i}}\left\|\sum x\right\rangle$
Bias : b	Displacement Operator :$\hat{\mathcal{D}}$	$\hat{\mathcal{D}}(\alpha)\left\|x\right\rangle = \left\|x+\alpha\right\rangle$
Activation Function : ϕ	Non-Gaussian transformation : Φ	$\phi\|x\rangle = \|\Phi(x)\rangle$



(a) Standard CVQNN architecture (b) Modified CVQNN architecture

Fig. 2. Three qumode CVQNN architectures

5 Experimental Results

Both the QBMLP and CVQNN have been applied to study the rise of corona virus cases in India and USA. The problem is same as multi dimensional regression analysis done commonly using classical neural nets. The networks have been trained to map non linear functions on basis of given data points and their performance has been compared on basis of two tailed t-test. A 3-3-1 architecture has been designed for QBP (i.e. 3 input neurons, 3 hidden neurons and 1 output neuron) and for the CVQNN, a 3 qumode quantum circuit has been constructed with only one hidden layer shown in Figure 2 (a). The output of the first qumode is finally measured using homodyne measurement in position eigenstates. The four parameters used from the data are Confirmed cases, Number of Deaths and Recoveries and number of days passed. Three of these features have been used to predict confirmed cases and the number of deaths since this information is quite

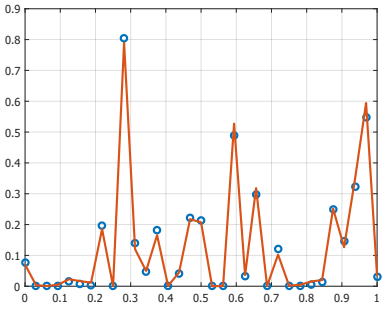
important with regards to public safety and social distancing standards. The activation function used is *tanh* as opposed to the more commonly used sigmoid function for quantum inspired model. The learning rate has been defined to be 0.001 for 15000 iterations(I_{QBMLP}). Since a nested for loop has been used in the algorithm of QBMLP, it's time complexity is given by $\mathcal{O}(Len(y) * I_{QBMLP})$, where Len(y) is specified by the length of output column. For CVQNN, the learning rate has been chosen as 0.1 (using stochastic gradient descent) with 200 iterations(I_{CVQNN}). The Strawberry fields simulator which is used to perform the CV computation, utilizes time/memory as $\mathcal{O}(C^W * I_{CVQNN})$ where C is the cutoff dimension (a hyper parameter, equal to 10 for our case) and W is the number of wires (or qumodes) involved in the computation which were 3 in this case, as their were 3 features. The caveat is, if we choose cutoff dimension to be low in order to speed up the computation, some operations like Displacement, Squeezing, Kerr operations push the quantum states out of the defined space giving low fidelity. Hence, the standard layer architecture has been modified to a novel version to speed up the computation as shown in Figure 2 (b). As we can see from the algorithm the CVQNN has better time complexity than our QBMLP model, but owing to sluggish nature of Strawberry fields simulator the former takes longer to execute in real time than the latter. The prediction results and the cost function decay for QBMLP and CVQNN for the test data set are shown in Figures 3 and 4, respectively.

Table 2. Two tailed t-test for QBMLP and CVQNN

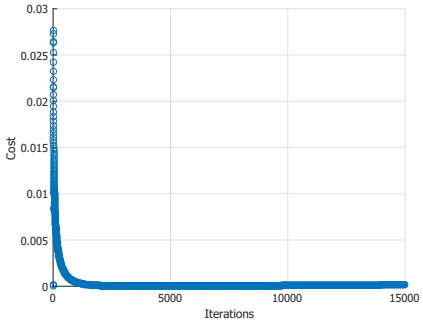
Data Set/Model	QBMLP		CVQNN	
	Statistic	p value	Statistic	p value
India - Death Prediction	0.0572	0.9545	0.01998	0.98411
India - Confirmed Prediction	-0.0189	0.9849	0.06372	0.94939
USA - Death Prediction	-0.02261	0.9820	-0.01964	0.98438
USA - Confirmed Prediction	0.0612	0.95132	0.08899	0.92936

6 Discussions and Conclusion

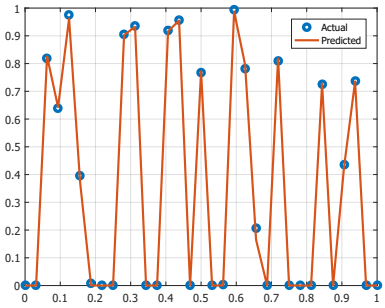
In this study the ability of quantum neural networks to perform regressional tasks has been investigated, specifically in predicting the rise in COVID-19 cases. The operational features used are the number of days passed, number of confirmed, recovered and death cases. The performance of a Quantum Backpropagation Multilayer Perceptron and a Continuos Variable Quantum Neural Network has been compared for the above mentioned task on the basis of a two tailed statistical t-test and it is found to be almost similar. While the QBMLP model performs better for checking the rise of confirmed cases in both the countries, the CVQNN model outperforms the former for calculating the rise in the number of deaths. Although the difference in p values for both the models is not that significant, it provides an interesting insight into the nature of both the models.



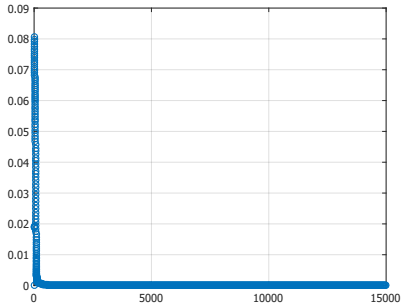
(a) India : Death cases



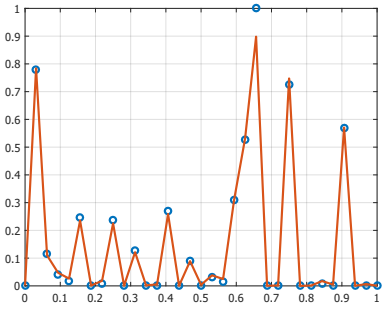
(e) Cost function decay (a)



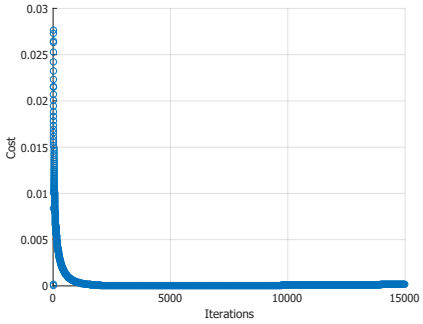
(a) USA : Death cases



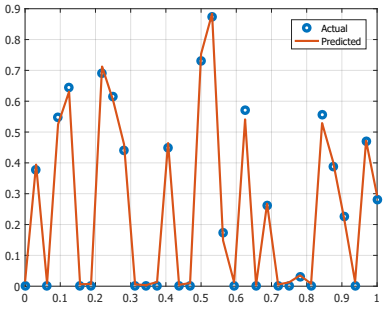
(f) Cost function decay (b)



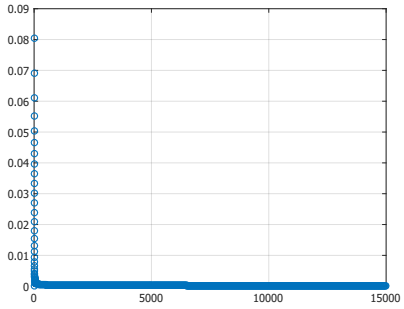
(a) India : Confirmed cases



(g) Cost function decay (c)

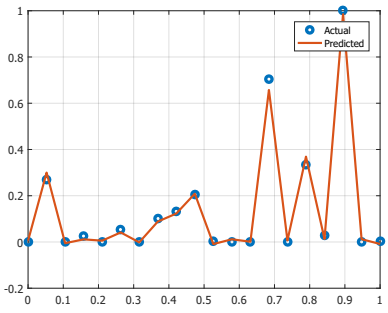


(a) USA : Confirmed cases

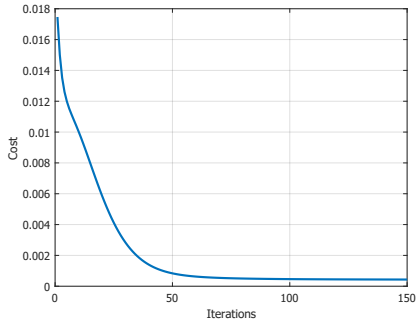


(h) Cost function decay (d)

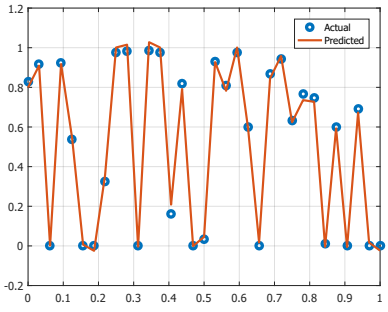
Fig. 3. Results and cost function decays of QBMLP for number of deaths and confirmed cases for COVID-19



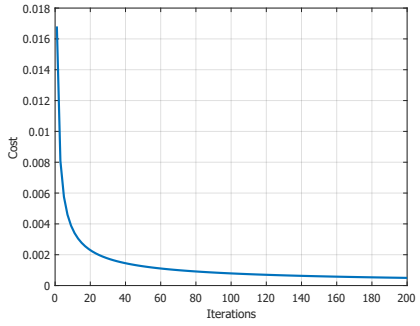
(a) India : Death cases



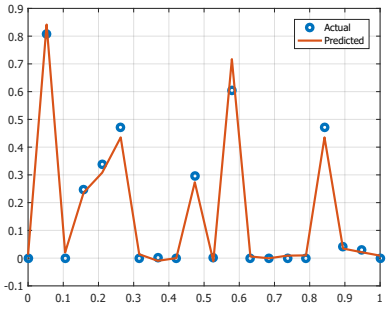
(e) Cost function decay (a)



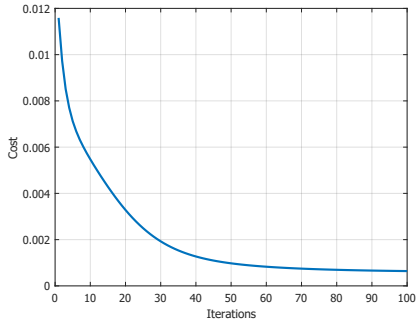
(b) USA : Death cases



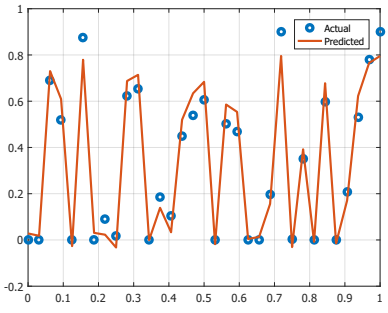
(f) Cost function decay (b)



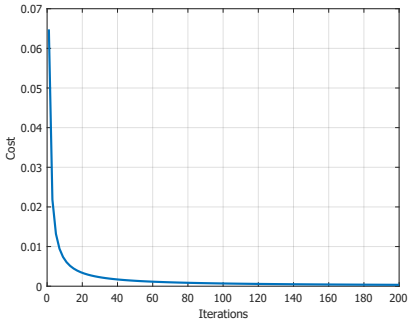
(c) India : Confirmed cases



(g) Cost function decay (c)



(d) USA : Confirmed cases



(h) Cost function decay (d)

Fig. 4. Results and cost function decays of CVQNN for number of deaths and confirmed cases for COVID-19

References

1. A. Darwish, Y. Rahhal, A. Jafar: A comparative study on predicting influenza outbreaks using different feature spaces: application of influenza-like illness data from early warning alert and response system in syria. BMC Research Notes, 13 (2020).
2. Y.-Y. Zheng et al.: COVID-19 and the cardiovascular system. Nature Reviews Cardiology, 17, pp. 259-260 (2020).
3. Y. Bai et al.: Presumed asymptomatic carrier transmission of COVID-19. JAMA, 323, pp. 1406 (2020).
4. F. Koike, N. Morimoto: Supervised forecasting of the range expansion of novel non-indigenous organisms. Alien pest organisms and the 2009 h1n1 flu pandemic, Global Ecology and Biogeography, 27, pp. 991-1000 (2018).
5. A. Steane: Quantum computing. Reports on Progress in Physics, 61, pp. 117-173 (1998).
6. T. Monz, et al.: Realization of a scalable shor algorithm. Science 351, 6277, pp. 1068-1070 (2016).
7. K. Huarng, T. H.-K. Yu: The application of neural networks to forecast fuzzy time series. Physica A: Statistical Mechanics and its Applications, 363, pp. 481-491 (2006).
8. Mitrpanont, Jarernsri L., and Ananta Srisuphab : The realization of quantum complex-valued backpropagation neural network in pattern recognition problem. In : 9th International Conference on Neural Information Processing. ICONIP'02., vol. 1, pp. 462-466. IEEE, (2002).
9. N. Killoran et al.: Continuous-variable quantum neural networks. Physical Review Research, 1 (2019).
10. D. N. Diep, K. Nagata, T. Nakamura: Nonparametric regression quantum neural networks.arXiv preprint arXiv:2002.02818 (2020).
11. S. Bhattacharyya, S. Bhattacharjee, N. K. Mondal: A quantum backpropagation multilayer perceptron (QBMLP) for predicting iron adsorption capacity of calcareous soil from aqueous solution. Applied Soft Computing, 27, pp. 299-312 (2015).
12. V. Bergholm et al.: Pennylane: Automatic differentiation of hybrid quantum-classical computations. arXiv preprint arXiv:1811.04968 (2020).
13. N. Killoran et al.: Strawberry fields: A software platform for photonic quantum computing. Quantum, 3, pp. 129 (2019).
14. Y. Liu et al.: J. Wu: Variational quantum circuits for quantum state tomography. Physical Review A, 101 (2020).
15. The Humanitarian Data Exchange. https://data.humdata.org/dataset/novel-coronavirus-2019-ncov-cases

View publication stats

Proceedings of the Second International Conference on Inventive Research in Computing Applications (ICIRCA-2020)
IEEE Xplore Part Number: CFP20N67-ART; ISBN: 978-1-7281-5374-2

Designing and Performance Analysis of 15KWP Grid Connection Photovoltaic System Using Pvsyst Software

Akash Chauhan

Department of applied Physics, Delhi Technological University (Formerly Delhi College of Engineering, University of Delhi) Bawana Road, Delhi-110042, India

Mayank Sharma

Department of applied Physics, Delhi Technological University (Formerly Delhi College of Engineering, University of Delhi) Bawana Road, Delhi-110042, India

Sarita Baghel

Department of applied Physics, Delhi Technological University (Formerly Delhi College of Engineering, University of Delhi) Bawana Road, Delhi-110042, India

Abstract - The most widely used photovoltaic systems are comprised of Grid Connection due to their reliability and efficiency. The concern now is to have a system, which can provide a suitable power with optimum conditions and provide significant performance over time. The below study is of such Grid Connected Photovoltaic array system which is designed to generate a nominal power of 15kWp at STC and the design is simulated using PVsyst software. The study focuses on key points like system production, output power losses with Sankey diagram, performance ratio (PR) and output energy of the system to be located, scrutinizing the output of PV module, adjusting the field type, sketching curves of power distribution and temperature distribution. This study also proposes the concern to module connection in string design and has shown the output distribution of inverter power.

I. INTRODUCTION

Power is an important ingredient to meet the daily need of today's world. Population growth has the adverse effect on power consumption. The increasing population have made the demand of energy continue to rise and deteriorated the resources of energy. To encounter the rising demand of energy, where it mainly depends upon the fossil fuels like petroleum, coal and natural gas which contain high percentage of carbon. Due to restricted obtainability, low efficiency and disadvantageous climate consequences caused by carbon rich fossil fuels the researchers aimed towards the utilization of renewable energy resources [1]. Solar energy emerged as one of the leading power source after discovery of photovoltaic (PV) cell. Solar energy has now become the third most source of power generation in the world because it converts the solar radiation into electricity directly. Solar photovoltaic energy shares a 25% of contribution in energy production making it the rapidly evolving power resource. Solar panels are now being used in industries, institutions and households [2], [3]. The growing field of solar cell to become more efficient and produce more power is a separate research topic. The crucial challenge now is to consume the available solar energy source in the most optimum manner such that it results with great efficiency. The photovoltaic (PV) systems are the solar power generating systems which are installed to generate electricity from solar radiation. The Grid Connected photovoltaic (PV) system are more favourable systems as they are easy to install and don't need battery system. This system holds the 99% of the capacity as compared to stand alone system throughout the world. This system covers a good life span and has no storage losses.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 07,2020 at 07:31:00 UTC from IEEE Xplore. Restrictions apply.

However the high penetration of photons can have negative effects to the grid [4], [6]. They avail bidirectional net energy metering which can reduce the usage bills and allow vending electricity to the power generation boards. The Standalone photovoltaic system is mainly for the remote areas and requires energy storage. They have storage losses and are cost inefficient. [4], [5]. In this paper, a 15kWp nominal power with Grid Connected Photovoltaic system is analysed using PVSYST Software.

II. PVSYST SIMULATION

Among various PV simulation software which is used to analyse the performance of PV system with different functioning conditions, PVSYST software provides a quick operation for the simulation process. It avails a large meteo and component data for preliminary and project design of Grid connected, Stand alone, Pumping and DC Grid photovoltaic system along with the economic evaluation parameters. Some studies have explained the Grid Connected PV system with its energy yield and economic parameters using PVSYST simulation software [7], [13]. Some researchers have examined different of photovoltaic cell types such as Polycrystalline, monocrystalline and amorphous Silicon using PVSYST simulation software [8-9]. Others authors evaluated the system by adjustment of tilt angle or changing the supplier brands of photovoltaic modules and inverters [8-9].

III. SITE AND METEOROLOGICAL DATA SELECTION

For location and area identity, where it is required to install the photovoltaic power plant. PVSYST facilitates with a large Meteo data that can be linked to our desired geographical area from sources like Meteonorm, NASA-SSE and PVGIS TMY. In this study, an area with coordinates (37.42° N latitude and -5.9° W longitude) near Seville city located at the southern Spain is considered. The location faces minimum 9 hours of sunlight for 320 days per year and with maximum of 15 hours one day during summers [11]. The table that contain Monthly geographical conditions of Seville on the basis of Meteonorm data source from PVSYST software is shown in table 1. Figure1 shows the yearlong satellite based data of Global Horizontal Irradiation in the solar map of Seville, Spain [11], [12], [13].

Month	Global Horizontal irradiation (kWh/m^2)	Diffuse Horizontal irradiation (kWh/m^2)	Temperature (°Celsius)
January	78.0	26.10	11.41
February	91.4	35.12	13.11
March	145.7	50.02	16.61
April	160.6	76.22	18.31
May	205.3	81.05	21.71
June	215.9	78.18	26.81
July	235.1	64.95	28.31
August	201.6	70.05	28.31
September	161.2	54.78	24.91
October	119.8	44.03	20.51
November	74.9	36.54	14.91
December	66.9	27.61	12.21
	1746.3	644.67	19.80

Table 1: Monthly Geographical Conditions



Figure 1: Global Horizontal irradiation per year

IV. SCHEMATIC DIAGRAM OF GRID CONNECTED PHOTOVOLTAIC SYSTEM

The Grid Connected Photovoltaic system is shown in figure 2. In this software generated design the PV array is connected to the Inverter which is connected to the user along with the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 07,2020 at 07:31:00 UTC from IEEE Xplore. Restrictions apply.

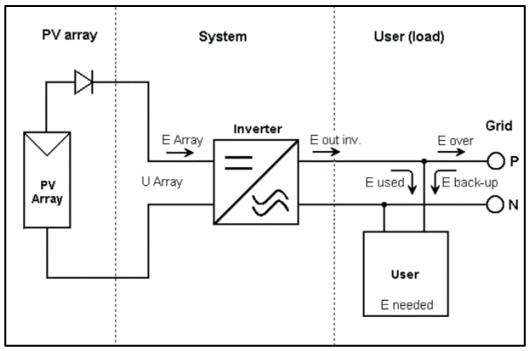
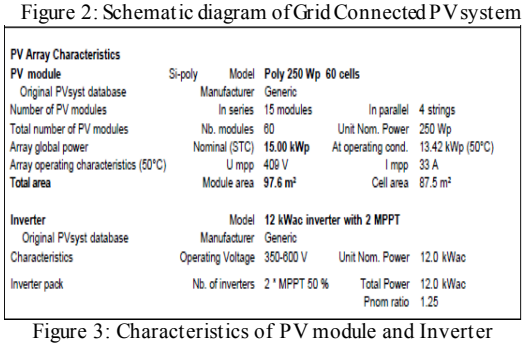


Figure 2: Schematic diagram of Grid Connected PV system



PV Array Characteristics					
PV module	Si-poly	Model	**Poly 250 Wp 60 cells**		
Original PVsyst database		Manufacturer	Generic		
Number of PV modules		In series	15 modules	In parallel	4 strings
Total number of PV modules		Nb. modules	60	Unit Nom. Power	250 Wp
Array global power		Nominal (STC)	**15.00 kWp**	At operating cond.	13.42 kWp (50°C)
Array operating characteristics (50°C)		U mpp	409 V	I mpp	33 A
Total area		Module area	**97.6 m²**	Cell area	87.5 m²
Inverter		Model	**12 kWac inverter with 2 MPPT**		
Original PVsyst database		Manufacturer	Generic		
Characteristics		Operating Voltage	350-600 V	Unit Nom. Power	12.0 kWac
Inverter pack		Nb. of inverters	2 * MPPT 50 %	Total Power	12.0 kWac
				Pnom ratio	1.25

Figure 3: Characteristics of PV module and Inverter

Grid. The design doesn't have any battery system as the electricity generated is vended to the Main grid [4], [5].This design depicts the way in which power generated and reaches to the user from the solar power plant.

V. SELECTION OF PV MODULE AND INVERTER

PV module selection mainly depends upon the type, power, efficiency and cost. These aspects vary with the type of module i.e. Mono-crystalline, Poly-crystalline and Amorphous [8], [9]. In this study various modules have been examined and from the available modules in PVSYST software, the 250Wp Poly-crystalline silicon PV module is selected with 60 solar cells manufactured by generic. The related specification is shown in table 2. For optimum operational condition of the selected PV module, the software recommends 15 modules in series and 4 strings in parallel. The efficiency variation over global incident radiation of selected PV module at various temperatures is shown in figure 4. The Direct Current (DC) current from the PV module is converted into Alternating Current (AC) with the help of an Inverter. From the examined inverters available in PVSYST software, a 12 kW inverter with 2 maximum power point tracker (MPPT) manufactured by Generic is selected for the simulation with the PV modules. The related specification is mentioned in table 3. The PV module and Inverter are the most important elements for designing of PV system hence detailed characteristics of the selection is shown in figure 3. The current vs Voltage curve giving the array voltage sizing is shown in figure 5 and the output energy vs array power curve is shown in figure 6.

Type	Poly-crystalline
Number of Modules	60
Maximum Power	250 Wp
Maximum Voltage	30 V
Open-Circuit Voltage	37.4
Maximum Current	8.33 A
Short-Circuit Current	8.63 A
Efficiency	17.25 %

Table 2: Manufacturer Specification and other measurement of PV module

Type	Grid inverter
Input Voltage (DC)	350-600 V
Input Current (DC)	38 A
Number of Phase	3 Phase
Number of inverters	2
Output Voltage (AC)	400
Efficiency	98 %

Table 3: Specification of Inverter

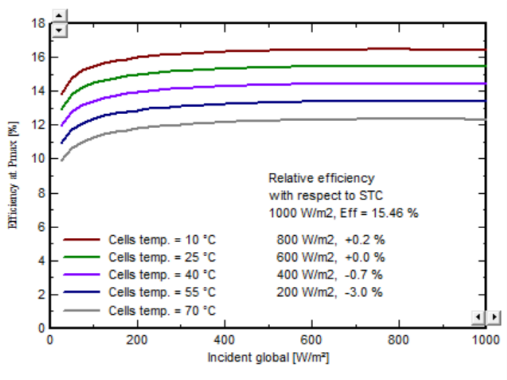


Figure 4: PV module- Generic poly 250Wp, 60 cells

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 07,2020 at 07:31:00 UTC from IEEE Xplore. Restrictions apply.

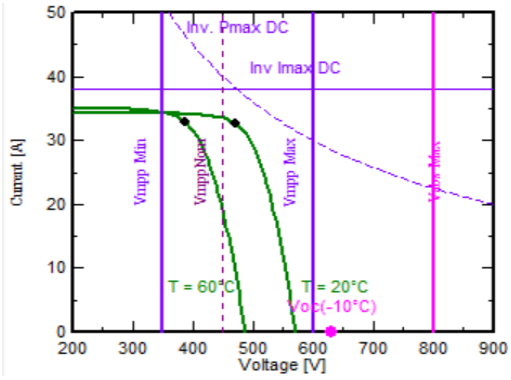


Figure 5: Array Voltage Sizing

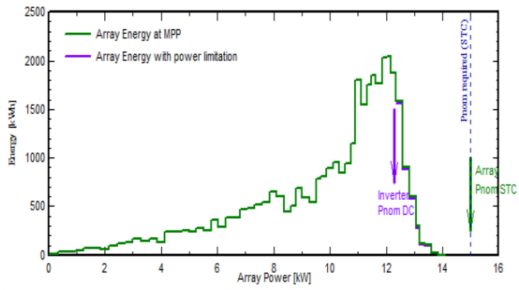


Figure 6: Power Sizing - Inverter output distribution

VI. ORIENTATION OF PV MODULE

The Maximum efficiency of PV module reaches up to 25% in the most favourable condition. These conditions depend on various factor which changes with location. Orientation is one of the factors which can increase the efficiency of a module. In this paper the solar trackers are installed at Seville, Spain located at Northwest of earth [11], [13]. This research work has selected the dual axis tracking modules, which moves according the direction of the sun. The tilt angle varies from 0° towards north to 80° towards south and the Azimuth angle varies from -120° towards east to 120° towards west [9] is shown in figure 7. In solar tracker the mutual shading is an important factor that can affect the performance of a PV module and its efficiency [13]. For the actual performance output of the system, linear mutual shading is introduced with the solar tracker to get near shading radiation loss. The compatible arrangement of panel with the orientation and shading construction is shown in figure 8.

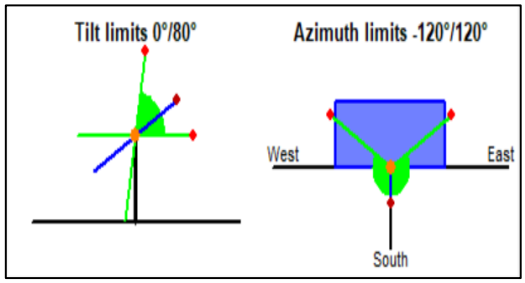


Figure 7: Tilt angle and Azimuth angle range

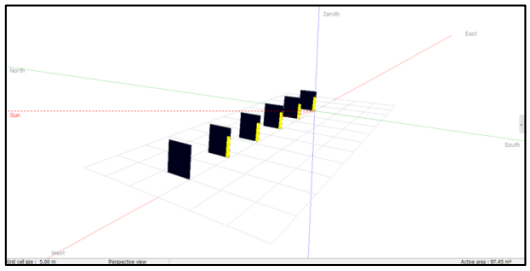


Figure 8: Solar panel arrangement with shading condition

VII. RESULTS

This section simulates the design and produce the results after considering the mentioned parameters and analysing the proposed system. The detail examined result of proposed 250Wp POLY CRYSTALLINE, Dual Axis Tracking PV module system simulated through PVSYST software shows the balance and main results, daily input/output energy diagram, normalized production, daily output energy, performance ratio and Sankey diagram. The Balance and main results evaluate the total power and performance ratio generated from 15kWp Poly Crystalline PV system is shown in figure 9.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 07,2020 at 07:31:00 UTC from IEEE Xplore. Restrictions apply.

	GlobHor kWh/m²	DiffHor kWh/m²	T_Amb °C	GlobInc kWh/m²	GlobEff kWh/m²	EArray kWh	E_Grid kWh	PR
January	78.0	26.10	11.41	167.9	165.3	2207	2156	0.856
February	91.4	35.12	13.11	167.5	162.5	2158	2107	0.839
March	145.7	50.02	16.61	238.4	225.6	2913	2845	0.796
April	160.6	76.22	18.31	225.1	210.6	2767	2703	0.801
May	205.3	81.05	21.71	284.4	269.1	3439	3361	0.788
June	215.9	78.18	26.81	296.8	283.3	3513	3432	0.771
July	235.1	64.95	28.31	330.7	315.5	3832	3743	0.755
August	201.6	70.05	28.31	293.2	275.7	3383	3308	0.752
September	161.2	54.78	24.91	250.7	236.6	2964	2897	0.770
October	119.8	44.03	20.51	205.1	196.9	2530	2472	0.803
November	74.9	36.53	14.91	139.5	135.9	1837	1794	0.858
December	66.9	27.61	12.21	147.2	144.5	1972	1927	0.873
Year	1756.3	644.67	19.80	2746.5	2621.5	33514	32746	0.795

Figure 9: Balance and main results

GlobHor: Horizontal global irradiation
DiffHor: Horizontal diffuse irradiation
T_Amb:Temperature ambient
GlobInc: Global incident in coll. Plane
GlobEff: Effective Global, corr. for IAM and shadings
E_Grid: Energy injected into grid
EArray: Effective energy at the output of the array
PR: Performance Ratio

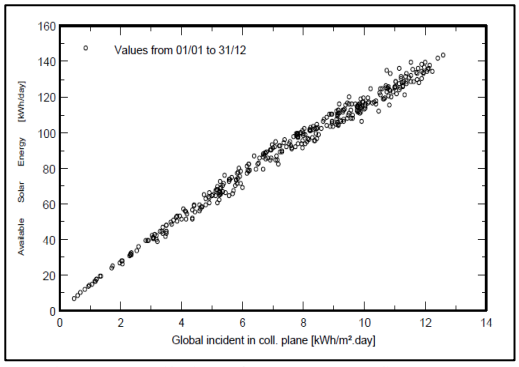


Figure 10: Daily input/output energy diagram

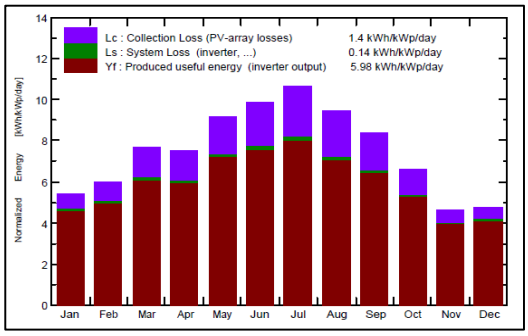


Figure 11: Normalized production

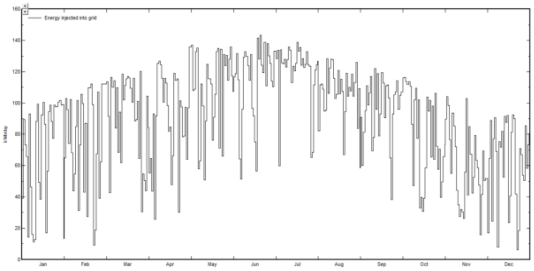


Figure 12: Daily system output energy

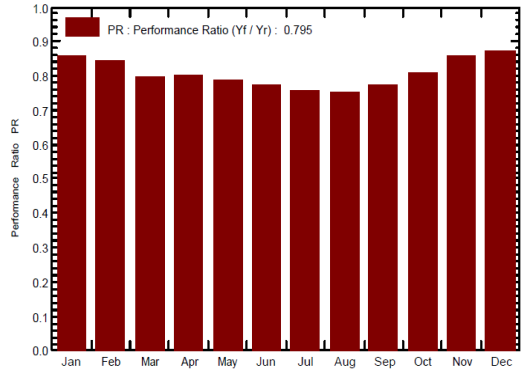


Figure 13: performance ratio

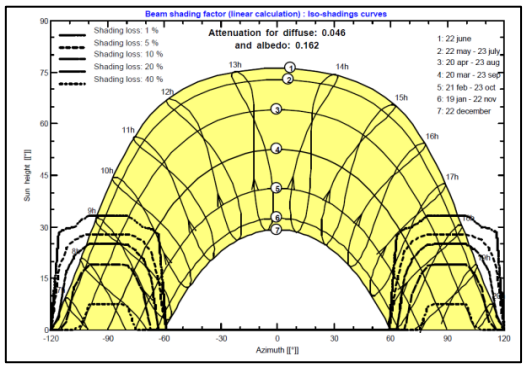


Figure 14: ISO-shading diagram

The Sankey diagram is the loss diagram of Photovoltaic system which describes the flow of power from irradiation to the output power injected to the grid after suffering various losses. Figure 15 shows the total loss is the difference between the Array nominal energy and energy injected to the grid which is nearly equal to 7.28 MWh.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 07,2020 at 07:31:00 UTC from IEEE Xplore. Restrictions apply.

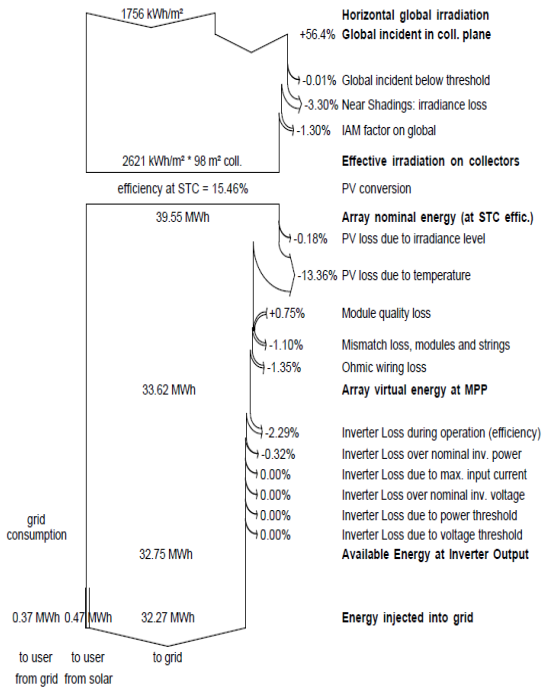


Figure 15: Sankey diagram over the year

VIII. CONCLUSION

This paper has studied a 15Wp connected PV solar power system. The installation area is located at Seville in Spain with latitude 37.42° North and -5.9° West. A Generic manufactured 250Wp, 60 cells Poly-Crystalline Dual axis PV module and Generic manufactured 12kWac inverter were selected for the system design. The solar tracker modules were also subjected to the linear shading construction to obtain the actual outputs. The simulation resulted with a performance ratio of 79.48% and a yearly production power of 32.272 MWh into the grid with a radiation loss due to linear shading of 3.3% is achieved. The injected energy into the grid is minimum 1.759 MWh in November and maximum 3.699 MWh in July is produced. The highest performance ratio 87.3% is achieved in the month of December. It is observed that the Photovoltaic module perform with higher efficiency in low temperature. The PVSYST software has provided a facility of detailed results and report to study the design module layout and orientation planning along with the detailed examination of losses.

ACKNOWLEDGEMENT

The authors acknowledge the initiatives and support of department of applied physics, Delhi Technological University (Formerly Delhi College of Engineering) Delhi.

REFERENCES

[1], Usenobong F. Akpan, Godwin E. Akpan, "The Contribution of Energy Consumption to Climate Change: A Feasible Policy Direction"
[2], Mohd Rizwan Sirajuddin Shaikh, Santosh B. Waghmare, Suvarna Shankar Labade, Pooja Vittal Fuke, Anil Tekale, "A Review Paper on Electricity Generation from Solar Energy"
[3] Mingzhi Zhao, Zhizhang Liu and Mingjun Yu, "Data acquisition and analyzing of solar energy resource", " Information and Automation (ICIA), Harbin, p.p., 2205-2208, 20-23 June 2010.
[4]http://www.fsec.ucf.edu/en/consumer/solar_electricity/basics/types_of_pv.htm
©2007-2014 University of Central Florida. The Florida solar energy centre (FSEC) is a research institute of the University of Central Florida
[5] B. Pakkiraiah and G. Durga Sukumar "Review Article Research Survey on Various MPPT Performance Issues to Improve the Solar PV System Efficiency.
[6], J Sreedevi (Joint Director), Ashwin N (Senior Research Fellow), M Naini Raju (Senior Research Fellow), "A Study on Grid Connected PV system".
[7] C. P. Kandasamy, P. Prabu and K. Niruba, "Solar potential assessment using PVSYST software", International Conference on Green Computing, Communication and Conservation of Energy (ICGCE) 2013, pp. 667-672.
[8] Nallapaneni Manoj Kumara, M. Rohit Kumara, Mobi Mathew, P. Ruth Rejoicea, "Performance analysis of 100 kWp grid connected Si-poly photovoltaic system using PVsyst simulation tool".
[9] Ahmed Alnoosani , Mowffaq Oreijah , Mohammed Alhazmi , Yousef Samkari , Hosam Faqeha, "Design of 100MW Solar PV on-Grid Connected Power Plant Using (PVsyst) in Umm Al-Qura University".
[10] Jake Lomansoc , Ricard Jun Roloma , Kim Seer Paller, "Design and Simulation of 100 kWp Solar Photovoltaic (PV) Grid Connected Power Plant Using PVsyst".
[11]https://www.suncalc.org/#/37.42,-5.9,9/2020.04.15/12:02/1/0
[12]https://apps.solargis.com/prospect/map?show-registration=1&s=37.42,-5.9&c=37.471952,-6.006409,9&m=solargis-ghi&l=true
[13] B. Kailash Krishna Prasad, K. Prahlada Reddy, K. Rajesh, P. Viswanath Reddy, "Design and Simulation Analysis of 12.4 kWp Grid Connected Photovoltaic system by using PVSYST Software".

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 07,2020 at 07:31:00 UTC from IEEE Xplore. Restrictions apply.

Drag Reduction by Application of Aerodynamic Devices in a Race Car

Devang S. Nath*
Mechanical Engineering Department
Delhi Technological University, India
devangsnath98@gmail.com

Prashant Chandra Pujari
Mechanical Engineering Department
Delhi Technological University, India
prashantpujari05@gmail.com

Amit Jain
Mechanical Engineering Department
Delhi Technological University, India
ajain1607@gmail.com

Vikas Rastogi
Mechanical Engineering Department
Delhi Technological University, India
rastogivikas@yahoo.com

Abstract

In this era of fast-depleting natural resources, the hike in fuel prices is ever-growing. With stringent norms over environmental policies, the automotive manufacturers are on a voyage to produce efficient vehicles with lower emissions. High-speed cars are at a stake to provide uncompromised performance but having strict rules over emissions drives the companies to approach through a different route to keep the demands of performance intact. One of the most sought-after ways is to improve the aerodynamics of the vehicles. Drag force is one of the major setbacks when it comes to achieving high speeds when the vehicle is in motion. This research aims to examine the effects of different add on devices on the vehicle to reduce drag and make the vehicle aerodynamically streamlined. A more streamlined vehicle will be able to achieve high speeds and consequently, the fuel economy is also improved. The three-dimensional car model is developed in SOLIDWORKS v17. Computational Fluid Dynamics (CFD) is performed to understand the effects of these add on devices. CFD is carried out in the

Research Square
Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

ANSYS^{TM} 17.0 Fluent module. Drag Coefficient (C_D) and Drag Force is calculated and is compared in different cases.

Keywords: Aerodynamics, CFD, Drag Force, Drag Coefficient, High speed car

Introduction

Aerodynamics is the study of how moving objects interact with the air. How the body behaves when it comes in contact with the air determines the forces induced by the air flowing over and around the body. It is one of the most important factors affecting the performance of a race car [1]. Driving the car is like swimming through the endless ocean of air. Over the past few years, the degrading air quality and the shortage of natural resources primarily oil, have tremendous pressure on automotive manufacturers to come up with some feasible solutions to overcome this crisis. In earlier times, high-speed cars were only dependent upon horsepower of the engine to maintain the performance segment of the vehicle. But in recent trends, design engineers are adapting the concepts of aerodynamics to enhance the efficiency of the vehicle [2,3]. Fuel consumption due to aerodynamic drag consumes about half the vehicle's energy [4,5]. Thus, reducing the drag is one of the major approaches automotive manufacturers opt for. Shaping the body of the vehicle and inclusion of various add on devices contributes to optimization for low drag, which becomes an essential part of the design process. Drag Force predominantly depends upon the velocity, frontal area, and coefficient of drag of the body. It can be expressed as:

$$F_d = 0.5\, C_d\, \rho\, A\, V^2$$

Where F_d is the drag force; ρ is the density of fluid medium that is air; A is the frontal area of the body facing the fluid; V is the velocity of the body; C_d is the coefficient of drag of the body.

From the above expression, it can be seen that the drag force is in proportion to the square of the speed. This implies that the resistance due to air increases exponentially as the speed of the body increases [6]. Flow separation control is also a major interest in fundamental fluid dynamics and various engineering applications [4,7]. Flow Separation location determines the size of the wake area and the amount of aerodynamic drag is determined accordingly. When the air moving over the vehicle is separated at the rear end, it leaves a large low-pressure turbulent region behind the vehicle known as the wake. This wake contributes to the formation of pressure drag [8].

Numerous techniques have been explored to control the flow separation either by preventing it or by reducing its effects [4].

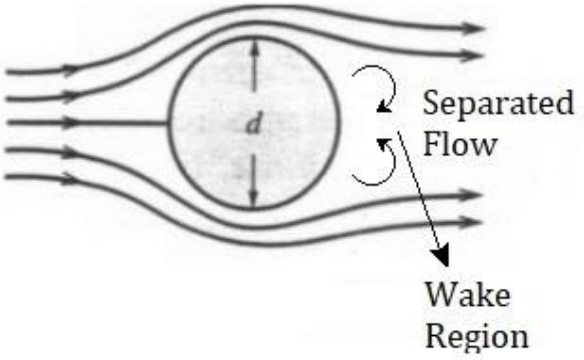


Fig. 1 Flow separation and formation of wake region

To achieve the optimized drag for the vehicle, the research is being carried out on this certain add on aerodynamic devices to reduce the resistance offered by wind and improve the efficiency of the vehicle [9]. In this research, the effects of various aerodynamic devices like the rear wing, spoiler, diffuser, and fins are examined and the change in the coefficient of drag is investigated.

Spoiler is one of the most widely used and important aerodynamic devices in the automotive domain. Its main purpose is to “spoil” the unwanted airflow and channeling the airflow in order which helps in reducing the drag. However, the actual use of spoiler is noticed at higher speeds approximately above 120 km/h. Commercial vehicles usually adopt it to increase the design appeal of the vehicle which provides little or no aerodynamic advantage. Thus, mostly high-performance vehicles adapt it to achieve higher speeds. The low-pressure zone behind the vehicle is reduced, thus less turbulence is created which subsequently leads to drag reduction.

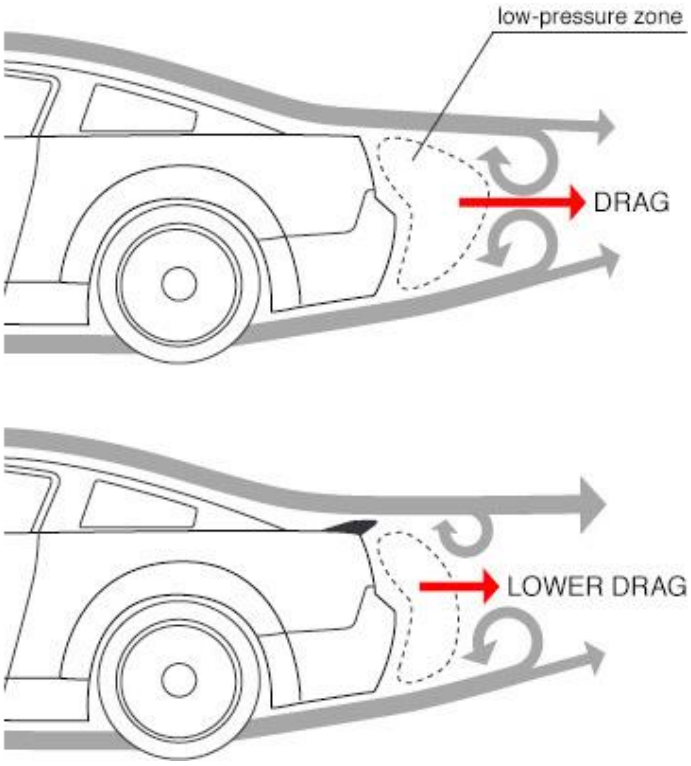


Fig. 2 Effect upon drag by using spoiler [10]

The wing is another essential aerodynamic device often used by race cars. A rear wing may look like a spoiler but is different in its functioning. It is shaped like a wing of an airplane turned upside down [6]. Its main objective is to provide sufficient downforce or negative lift so that the vehicle has increased traction and the vehicle doesn't lift off at higher speeds [11]. It also allows to corner faster and improves stability at high speeds [12]. But using a wing may add up the drag to the vehicle body. Thus, for any amount of lift gained, drag also increases [13]. It is generally regarded as a tradeoff between drag and lift.



Fig. 3 Wing at the rear of a car [14]

For the first time in the automotive industry, the application of fins at the rear part of the vehicle's body is witnessed by Swedish hyper-car manufacturer Koenigsegg Automotive AB. Their flagship model "Jesko Absolut" which has the least coefficient of drag in their lineup has fins instead of the wing as shown in Fig. 4. Fins are inspired by fighter jets to provide high-speed stability and to reduce aerodynamic drag.



Fig. 4 Koenigsegg Jesko Absolut [15]

The diffuser is one of the prominent aerodynamic devices found in Formula 1 cars. The wide versatility offered by diffusers has found its way down to the high-speed production vehicles. Diffusers are capable of reducing drag and increase downforce for

driving cars [16,17]. The role of the diffuser is to expand the flow from underneath the car to the rear, this in turn produces a pressure potential, which will accelerate the flow underneath the car resulting in reduced pressure [18]. The principle behind the working of diffusers is based upon Bernoulli's principle which states that "a slow-moving fluid will exert greater pressure than the fast-moving fluid". Thus, the role of the diffusers is to accelerate the flow of air beneath the car so that less pressure is exerted in comparison to the outer body flow. This serves in ejecting out the air from below the car. The diffuser then lightens this high-speed air down to normal speed and helps fill the area behind the car making the entire underbody a more robust downforce and importantly reducing the drag on the vehicle.

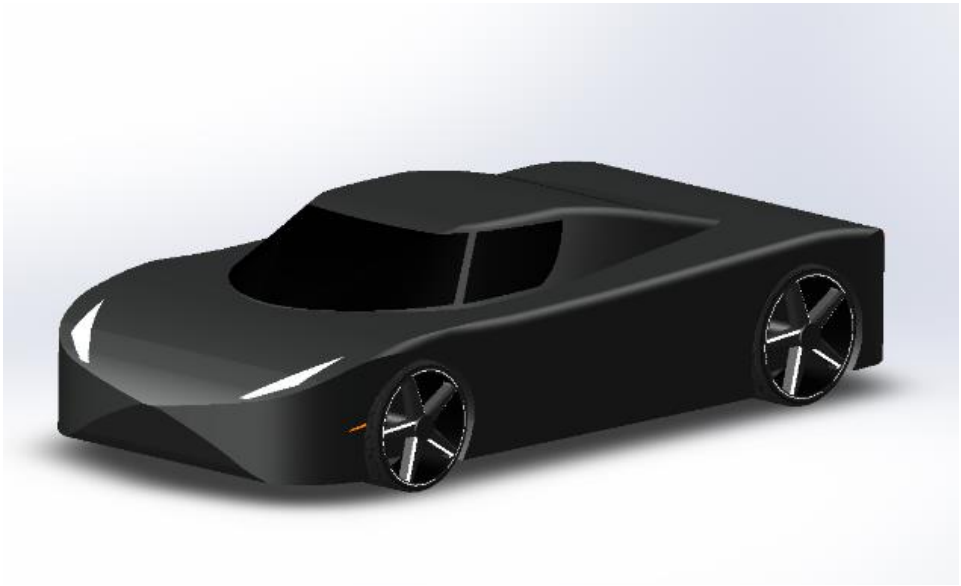


Fig. 5 Diffuser in a car [14]

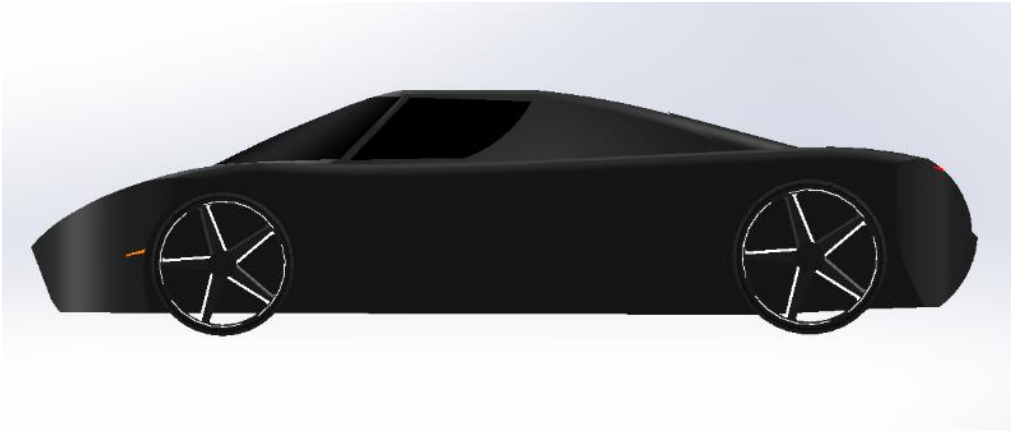
Geometric modelling

Baseline Model GT

The vehicle used for simulation is shown in Fig. 6. The three-dimensional car model was developed in SOLIDWORKS v17.0. The baseline model's length, width, height are 4230mm, 1996mm, 1089mm respectively. While the ground clearance is 92mm.



(a) Isometric view



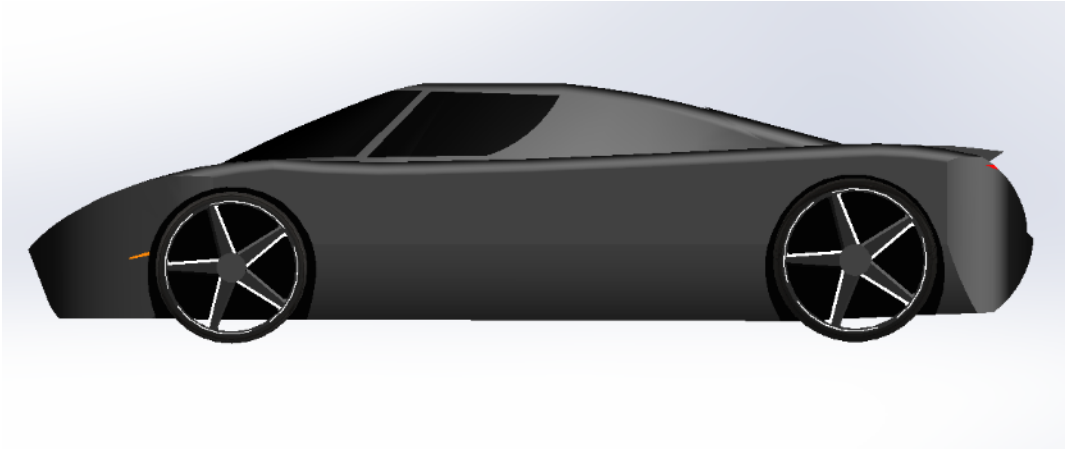
(b) Side view

Fig. 6 Geometric model of the car

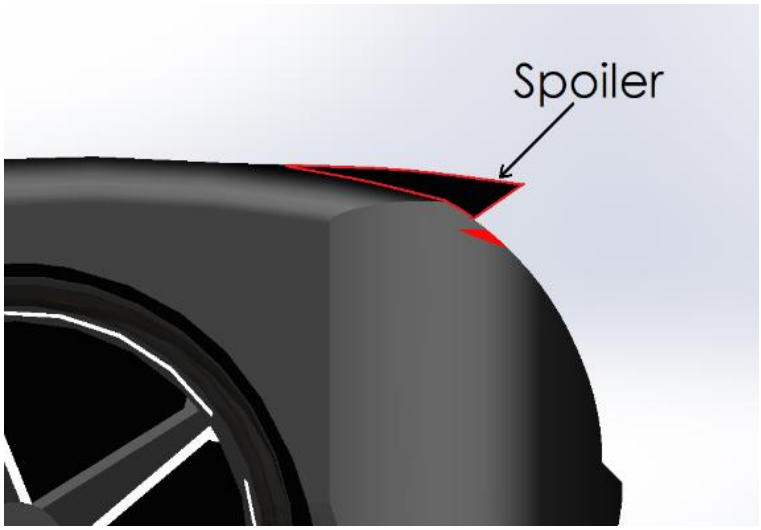
Seven different cases excluding baseline model GT with various add on devices namely, GT with spoiler, GT with wing, GT with diffuser, GT with fins, GT with both spoiler & diffuser, GT with both wing & diffuser and lastly GT with both fins & diffuser have been illustrated below.

GT with Spoiler

In the baseline GT model, a spoiler has been installed at the rear end of the trunk in Fig. 7 to provide a greater streamlined flow by delaying the flow separation in an attempt to reduce the overall drag coefficient of the car.



(a) Side view

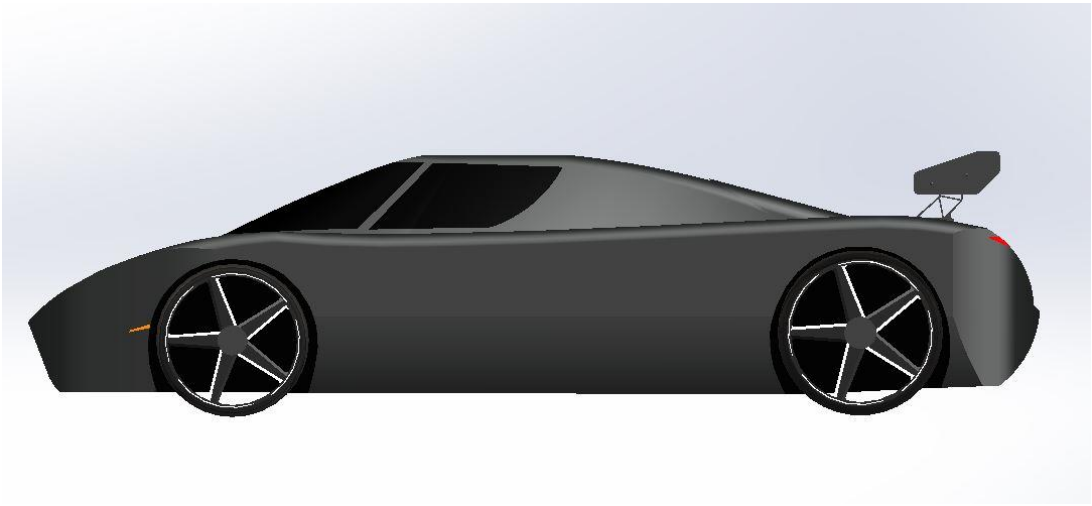


(b) Spoiler

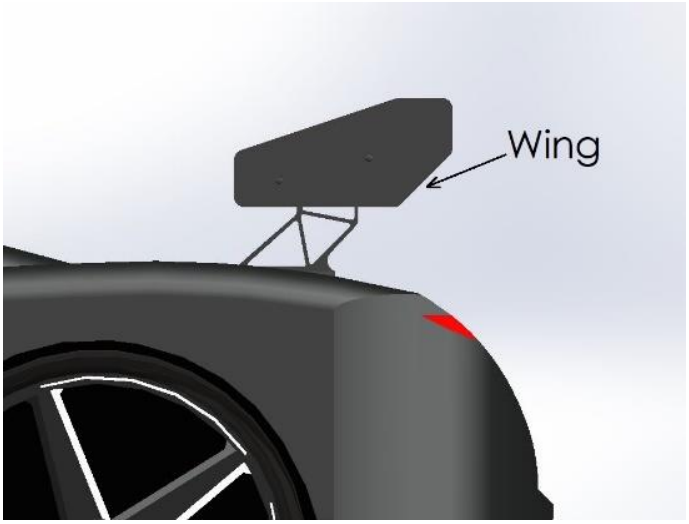
Fig. 7 GT with spoiler

GT with Wing

A wing 1800mm long span has been installed in the baseline model at the rear end of the trunk as shown in Fig. 8 with an angle of attack of 31.4 degrees. It is expected that the application of the rear wing will increase the downforce in the trade of drag force.



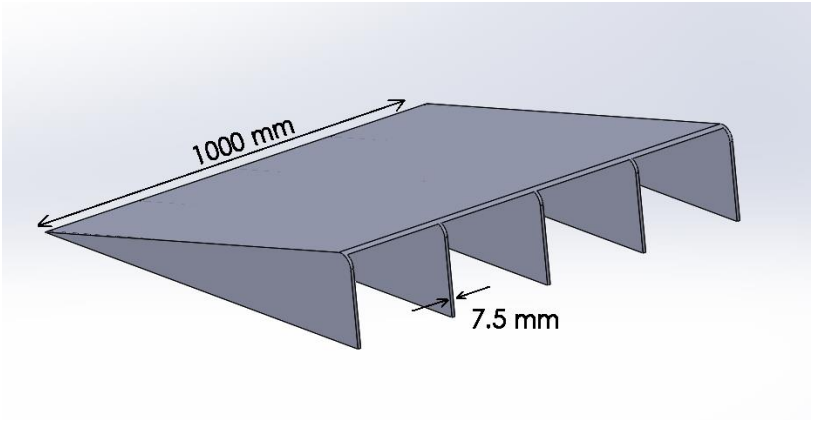
(a) Side view



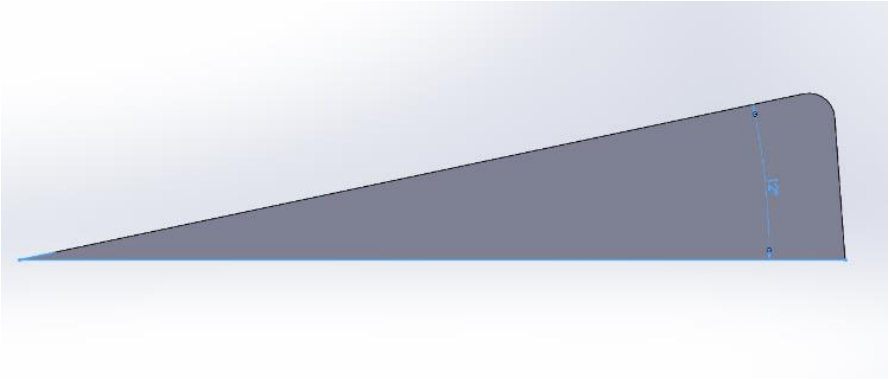
(b) Wing

Fig. 8 GT with wing

GT with Diffuser



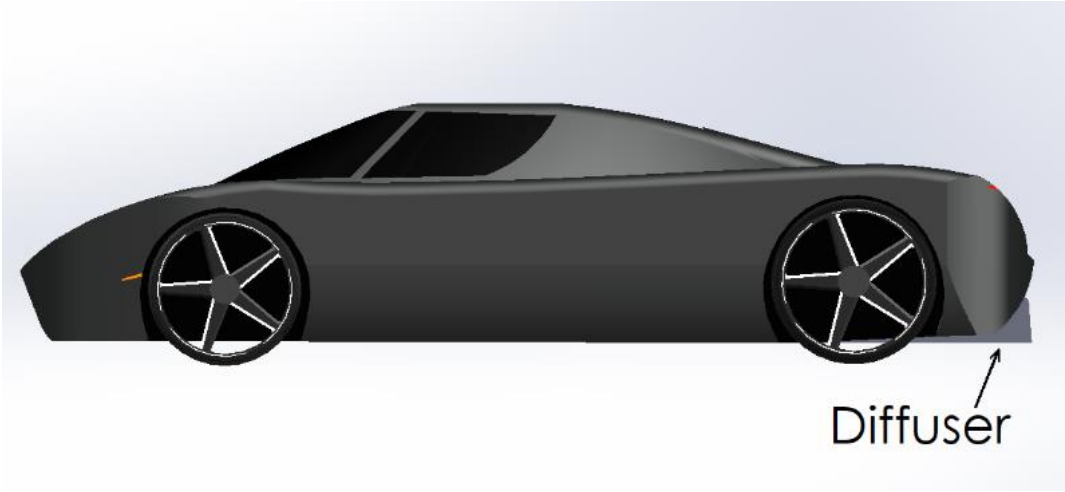
(a) Isometric view



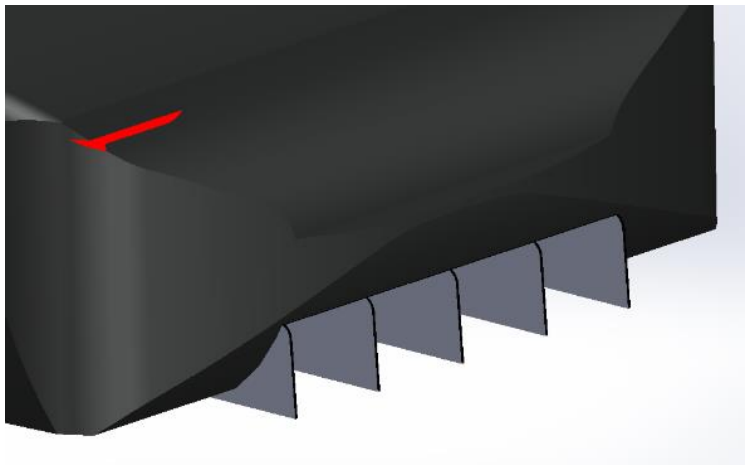
(b) Side view

Fig. 9 Diffuser model

A diffuser is developed of length 1000mm while its angle of inclination is 12 degrees and the width of each teeth is 7.50mm as shown in Fig. 9. A steeper angle of inclination may result in flow separation which will lead to an increase in drag. A lower angle of inclination will be less effective for the required purpose. Diffuser has been installed at the rear of the baseline GT as shown in Fig. 10.



(a) Side view

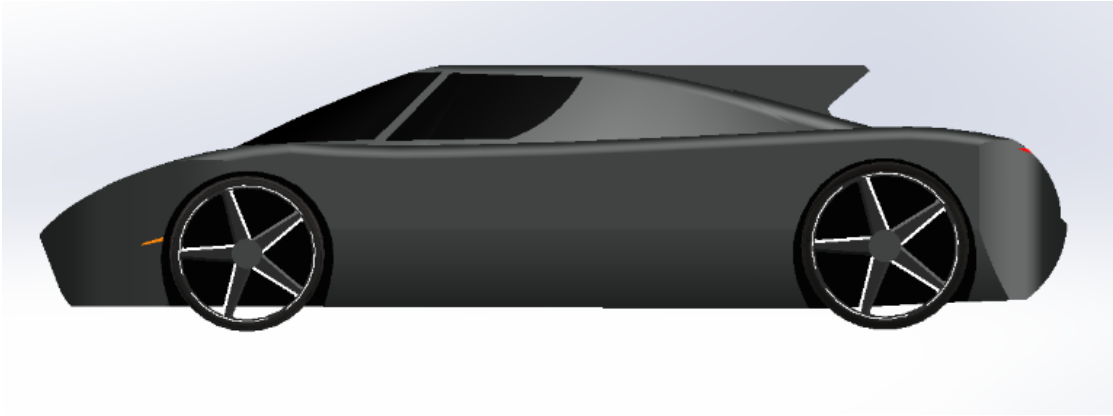


(b) Diffuser

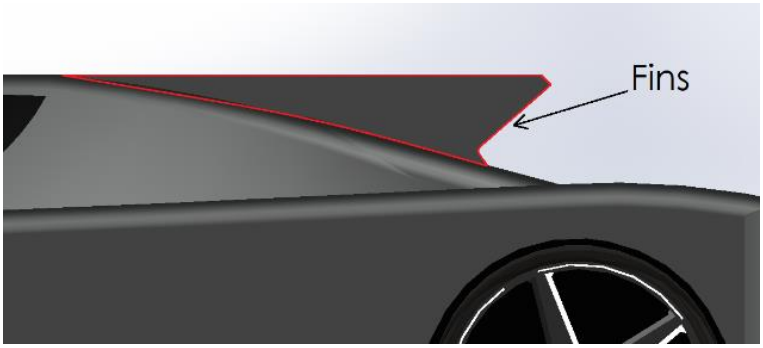
Fig. 10 GT with diffuser

GT with Fins

The Fins are installed at the rear part of the baseline GT, each of thickness 15mm as shown in Fig. 11 to provide overall stability at high speeds. It is expected that the overall drag coefficient of the vehicle will be decreased by the application of the fins.



(a) Side view

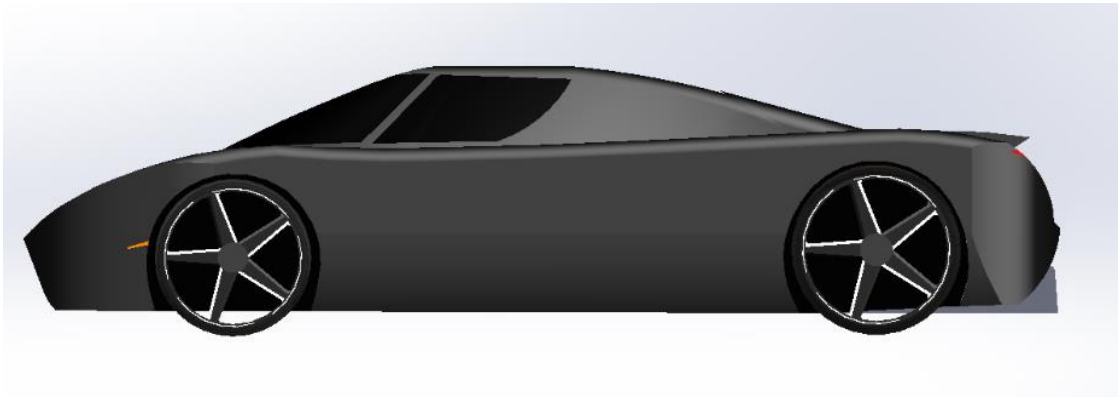


(b) Fins

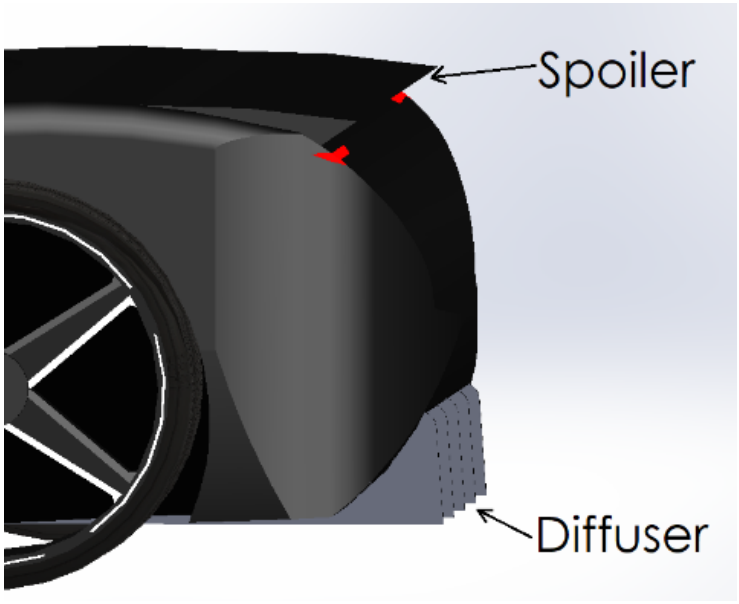
Fig. 11 GT with fins

GT with Spoiler and Diffuser

In the baseline GT model, both spoiler and diffuser are installed as shown in Fig. 12. It is expected that the combined aerodynamic devices will reduce the drag coefficient of the vehicle by a greater margin.



(a) Side view

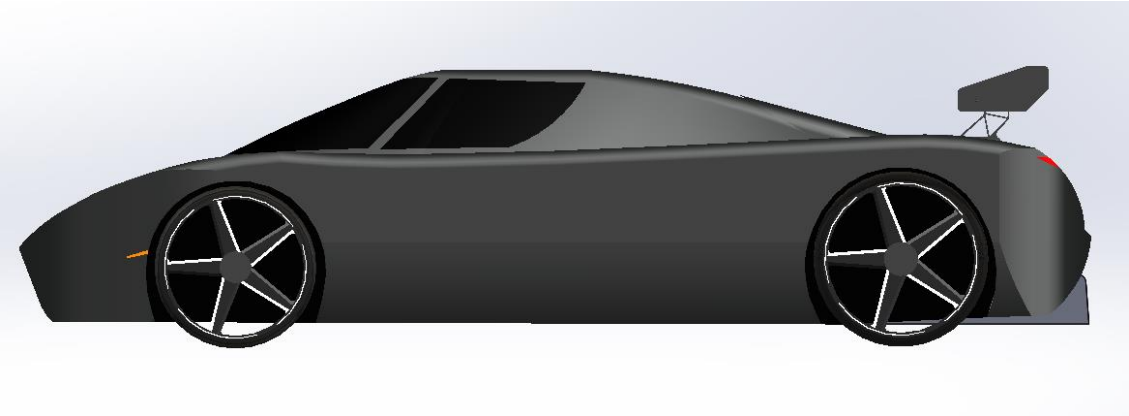


(b) Spoiler and diffuser

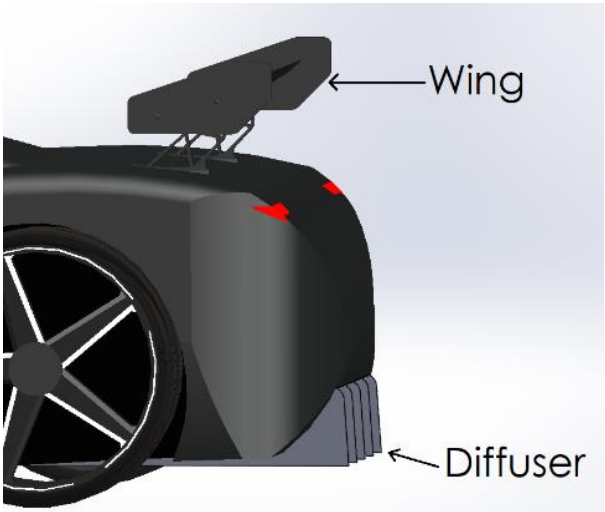
Fig. 12 GT with spoiler and diffuser

GT with Wing and Diffuser

In the baseline GT model, both wing and diffuser are installed as shown in Fig. 13. Application of both of these aerodynamic devices will provide downforce as well as a reduction in drag.



(a) Side view

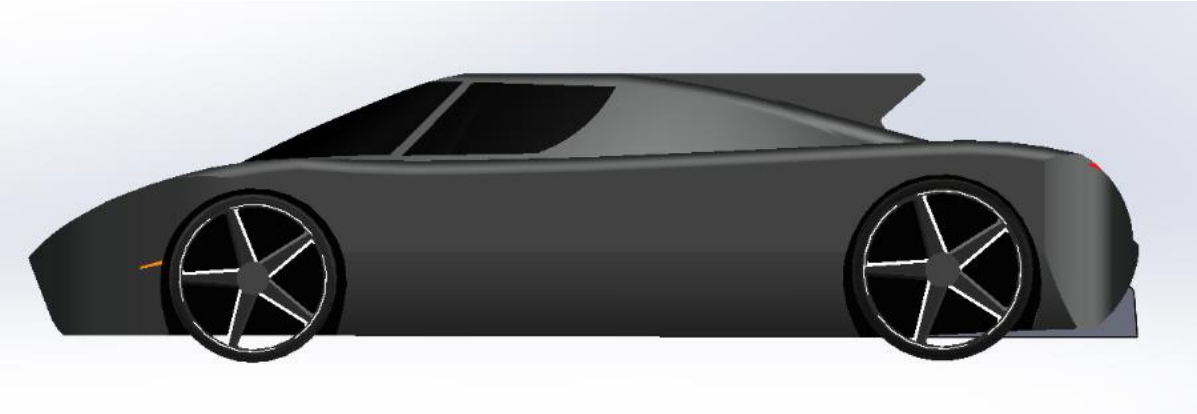


(b) Wing and diffuser

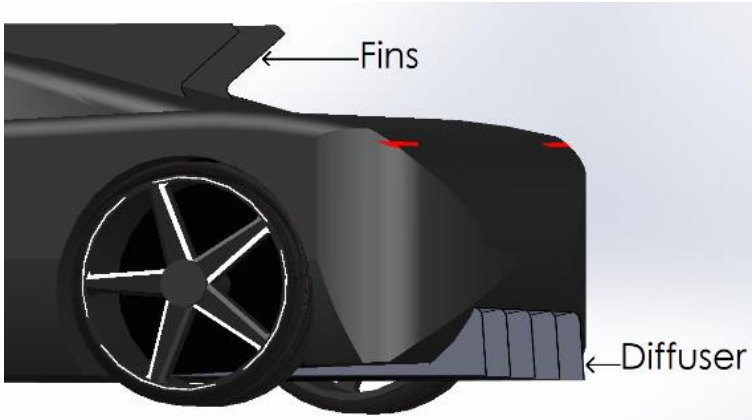
Fig. 13 GT with wing and diffuser

GT with Fins and Diffuser

In the baseline GT model, both fins and diffuser are installed as shown in Fig. 14. It is expected that the application of both of these aerodynamic devices would help in providing stability and reducing the overall drag coefficient.



(a) Side view



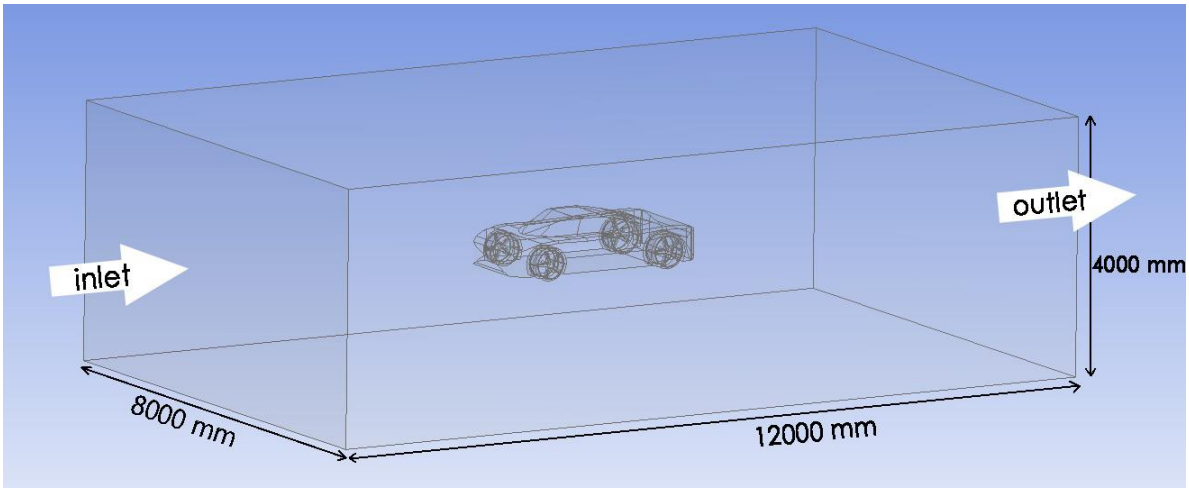
(b) Fins and diffuser

Fig. 14 GT with fins and diffuser

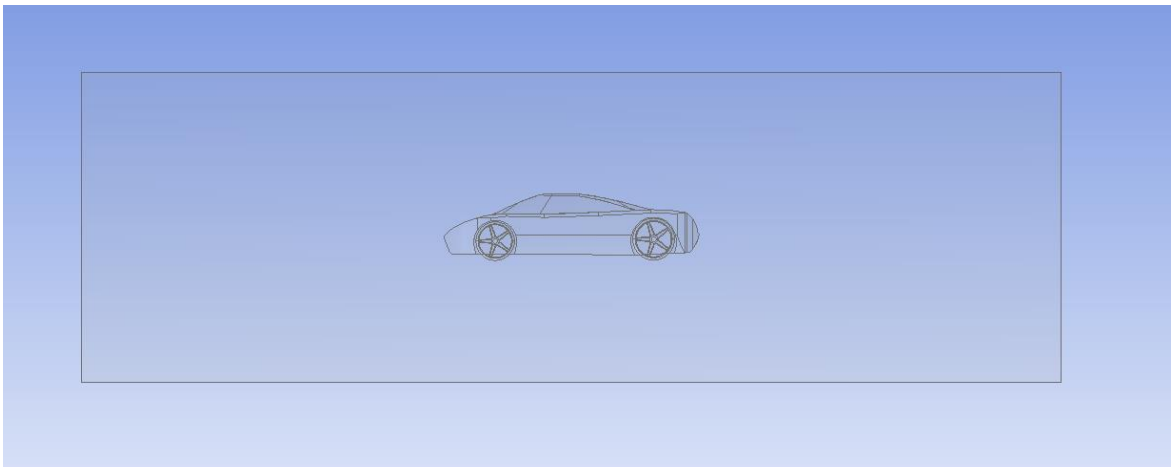
Simulation

Methodology

The three-dimensional car model was imported to ANSYS^{TM} workbench. Computational Fluid Dynamics (CFD) was carried out in the FLUENT module. In Design Modeler, an enclosure is developed of dimensions 12000 X 4000 X 8000 mm to form a virtual wind tunnel as shown in Fig. 15.



(a) Isometric View



(b) Side View

Fig. 15 Enclosure

Meshing

An appropriate mesh was developed using ANSYS^{TM} Mesh Tool. The mesh was observed to be coarser at the inner domain and finer at the contact region with the vehicle as shown in Fig. 16. A total of 227096 nodes and 1192687 elements were formed after the meshing was done.

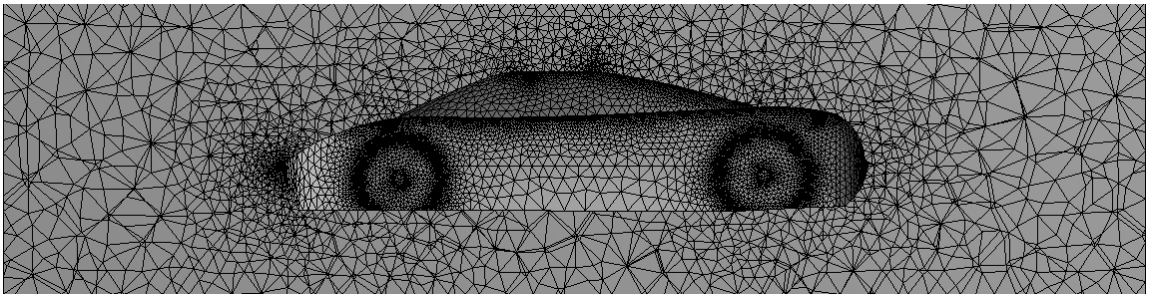


Fig. 16 Meshed model

Boundary Conditions

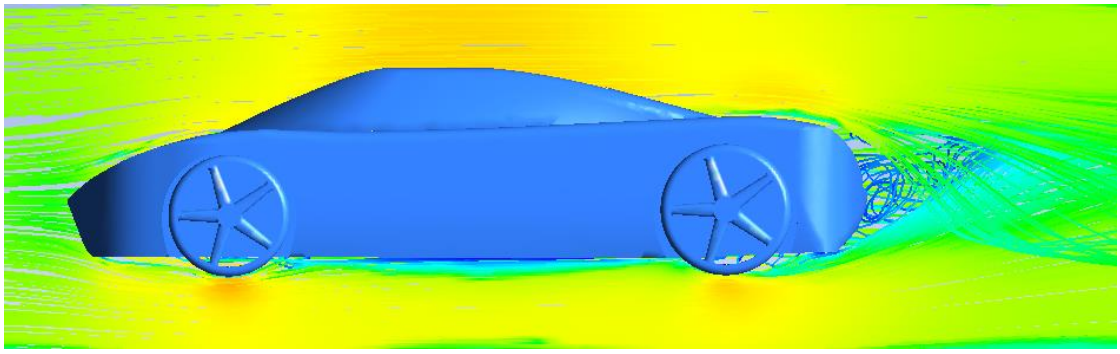
The numerical simulation was done in commercial code FLUENT [16]. Due to its stability and ease of convergence, the Standard k-epsilon model was selected as a turbulence model [18-19]. The coupled scheme was set as the iterative algorithm and the residual value was set to 0.001. The frontal surface area of the vehicle is 1.99820 m^2. A constant velocity boundary condition was selected for the inlet boundary. Since the majority of these devices make a difference at higher speeds, thus inlet velocity is kept at 150 kmph. The outlet was set to constant pressure conditions. Boundary Conditions used for the simulation are listed below in Table 1.

Table 1. Boundary conditions

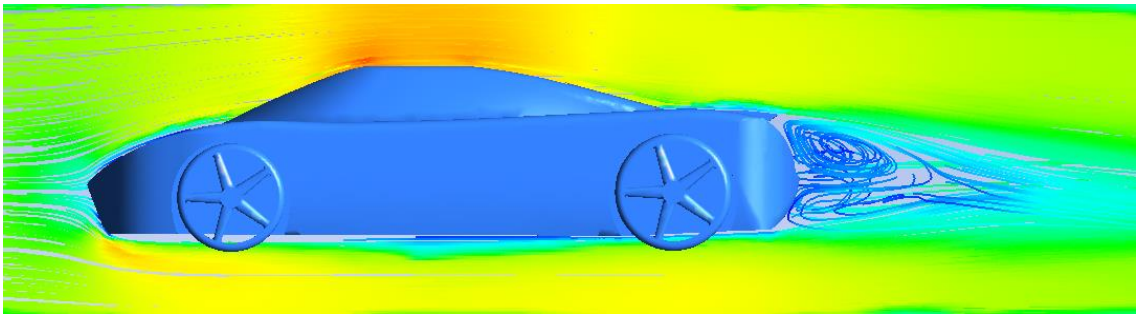
Region	Boundary condition
Inlet	Velocity inlet, v = 150 km/h
Outlet	Pressure outlet, reference pressure = 0 Pa
Vehicle Surface	No slip wall
Top and side	No slip wall
Ground	No slip wall
Temperature	288.16 K

Results and discussions

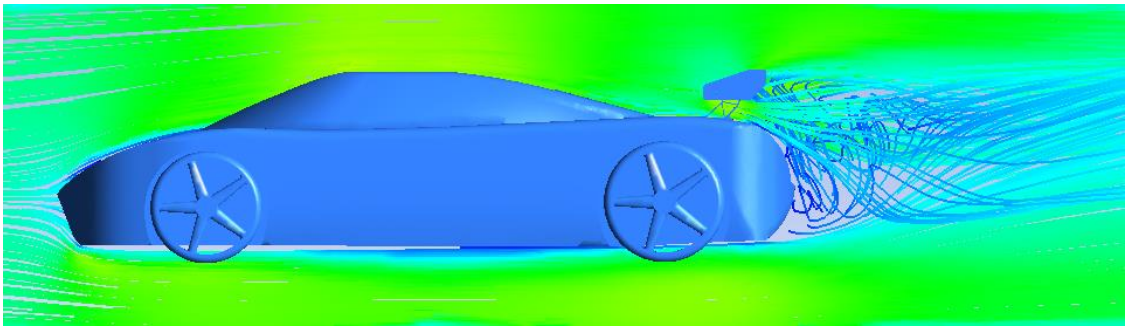
All the designed car models with different and combined aerodynamic devices are simulated in ANSYS^{TM} 17.0 Fluent. The results of the drag coefficients obtained are discussed below.



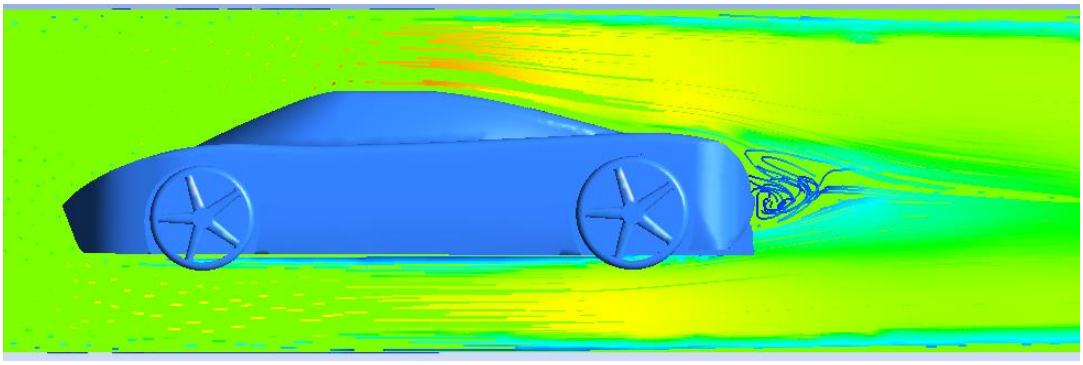
(a) GT baseline



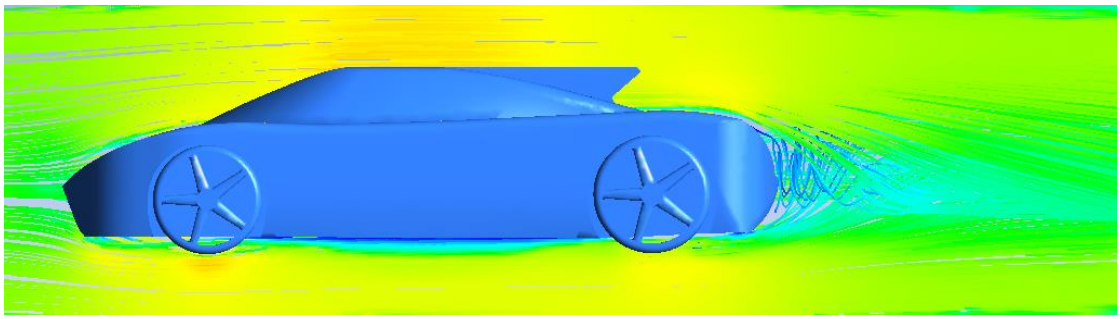
(b) GT with spoiler



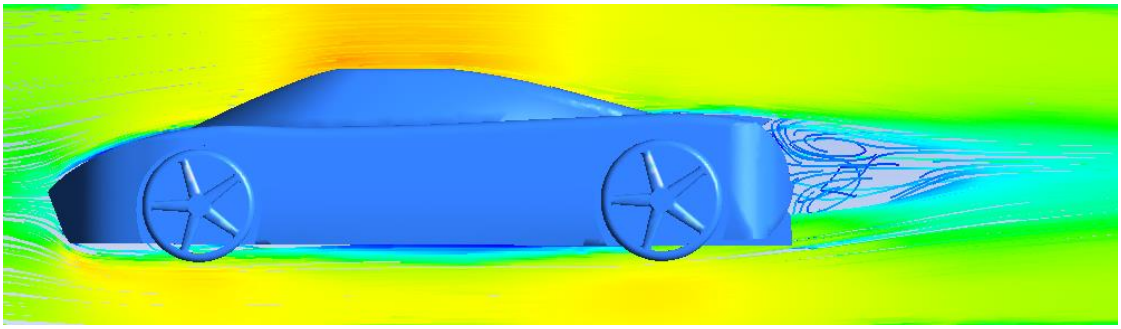
(c) GT with wing



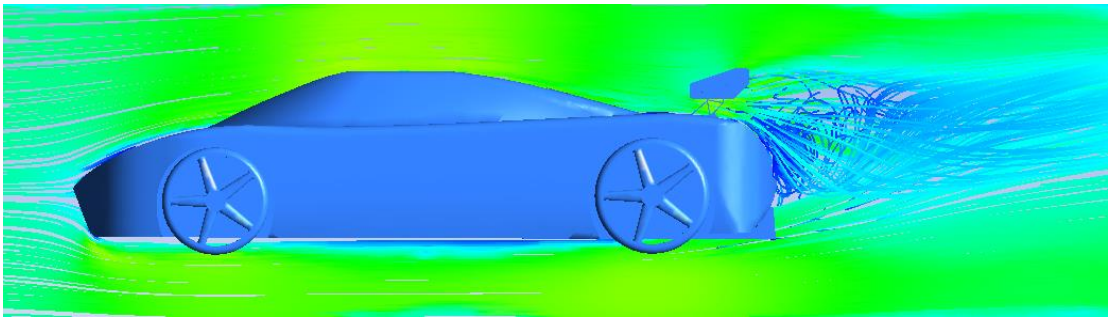
(d) GT with diffuser



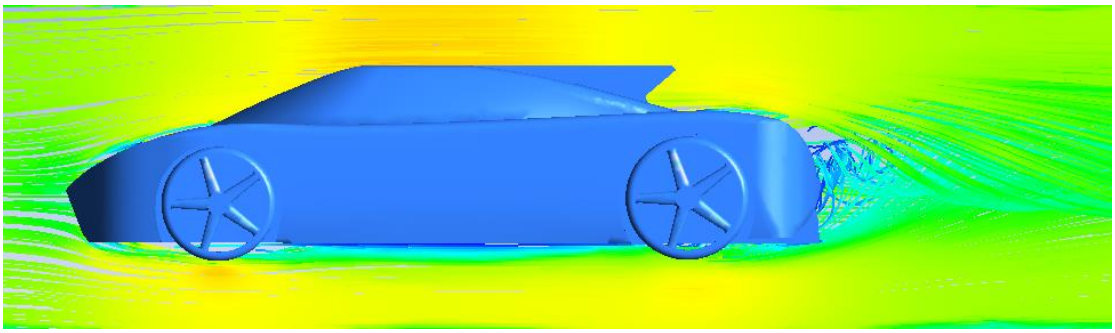
(e) GT with fins



(f) GT with spoiler and diffuser



(g) GT with wing and diffuser

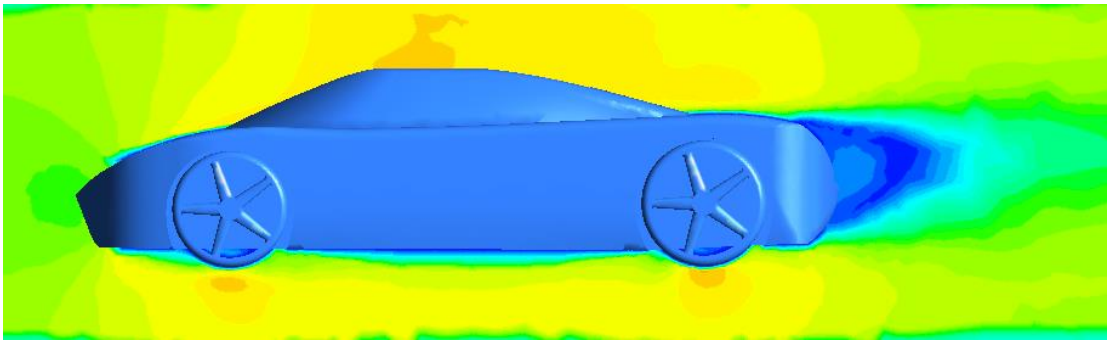


(h) GT with fins and diffuser

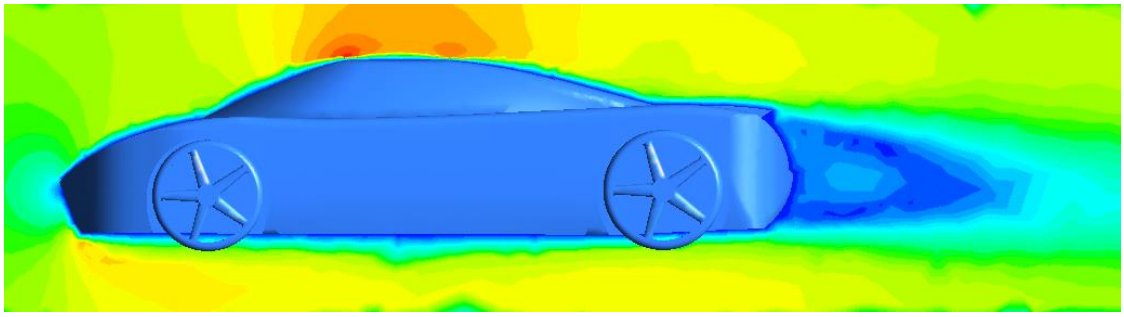
Fig. 17 Streamlines after vehicle for different cases

Figure 17 depicts the streamlines plot derived from the simulation for different cases. As seen from the streamlines plot above, minimum flow separation is favorable which subsequently leads to lesser turbulence. The amount of turbulence created behind the rear region of the car determines the magnitude of the drag force. In the case of GT with Wing Fig. 17 (c), maximum flow separation is observed which has led to maximum turbulence, thus maximum drag force. On the other hand, the application of the Spoiler Fig. 17 (b) on the baseline model reduced the turbulence at the back, thus drag force is also reduced considerably. Moreover, the addition of Diffuser Fig. 17 (d) to the baseline model accounted for streamlined flow, which reduced the drag force. In the combination of these two in the case of GT with Spoiler and Diffuser Fig. 17 (f), it is evident from the streamlines that the least turbulence is generated, which further reduced the drag coefficient minutely, thus the least drag force out of all the cases stated above.

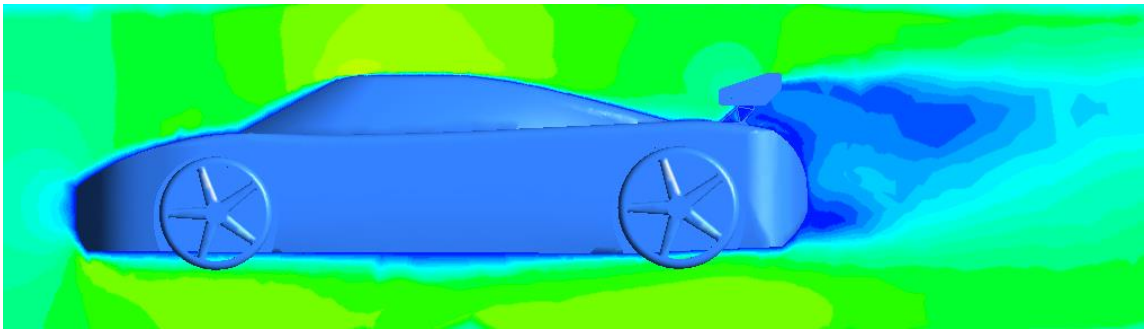
Figure 18 shows the velocity contour for all the cases. From velocity contours, the re-circulation zone is visualized behind the vehicle. The minimum the recirculation zone, the least turbulence created which subsequently leads to minimized drag.



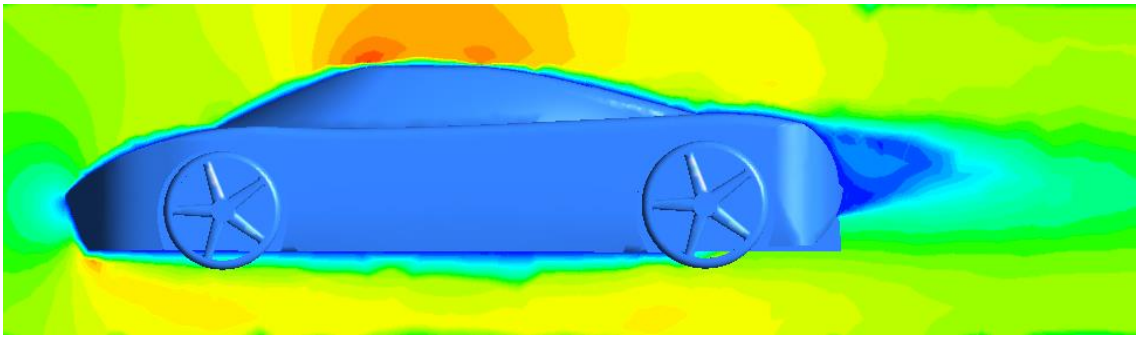
(a) GT baseline



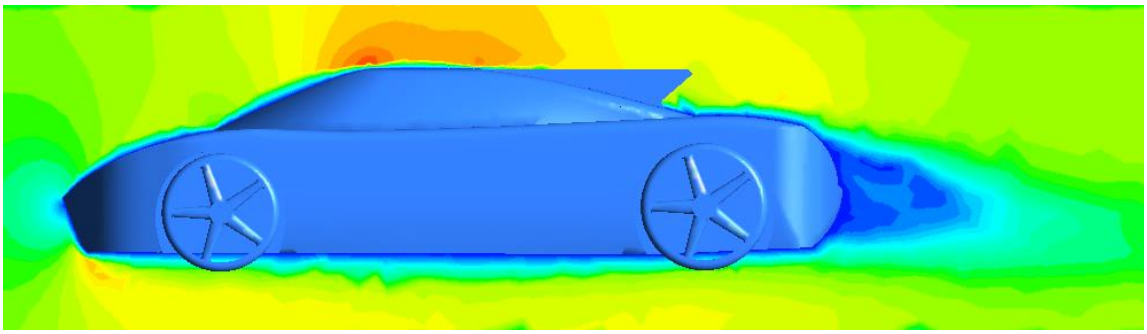
(b) GT with spoiler



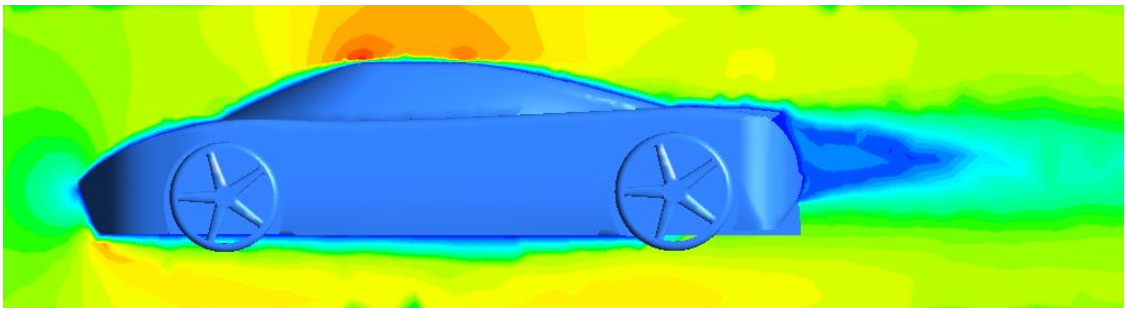
(c) GT with wing



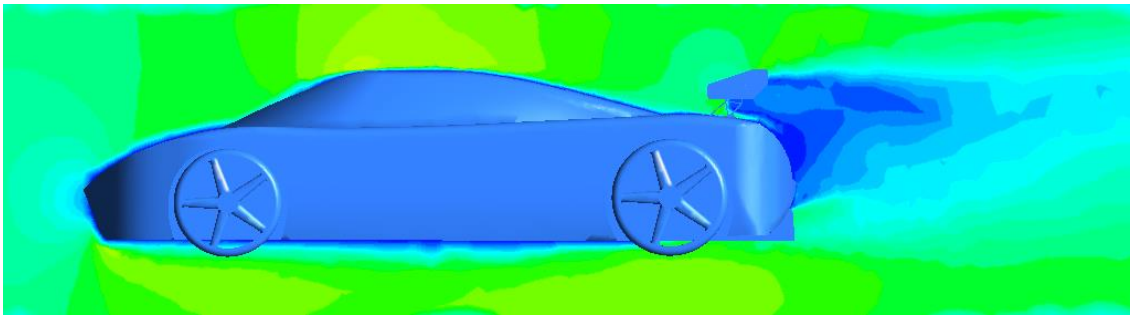
(d) GT with diffuser



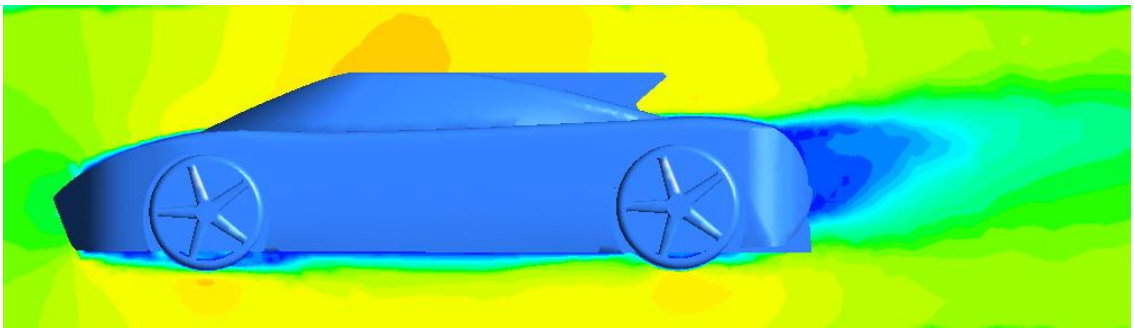
(e) GT with fins



(f) GT with spoiler and diffuser



(g) GT with wing and diffuser



(h) GT with fins and diffuser

Fig. 18 Velocity contours behind the vehicle for different cases

From velocity contours derived from the simulation, it is evident that in the case of GT with Wing Fig. 18 (c), there is a large re-circulation zone extending from the bottom of the trunk to the wing's edge. Thus, it has contributed to the maximum drag force. Moreover, the addition of a Spoiler Fig. 18 (b), the recirculation zone is reduced from the baseline model, which reduced the drag force considerably. On the application of Diffuser Fig. 18 (d), the recirculation zone is reduced, accounting for more teardrop shape with less flow separation, which reduced the drag force. Furthermore, the combination of these two in the case of GT with Spoiler and Diffuser Fig.18 (f), the recirculation zone is small, which has contributed to the even lesser drag out of all the cases. Moreover, as seen from the above images, it can be deduced that diffuser tends to decrease the recirculation zone to have a more teardrop shape. Thus, organizing the flow and eventually drag coefficient of the vehicle is reduced. Fig. 19 shows the comparison of the drag coefficient for different cases.

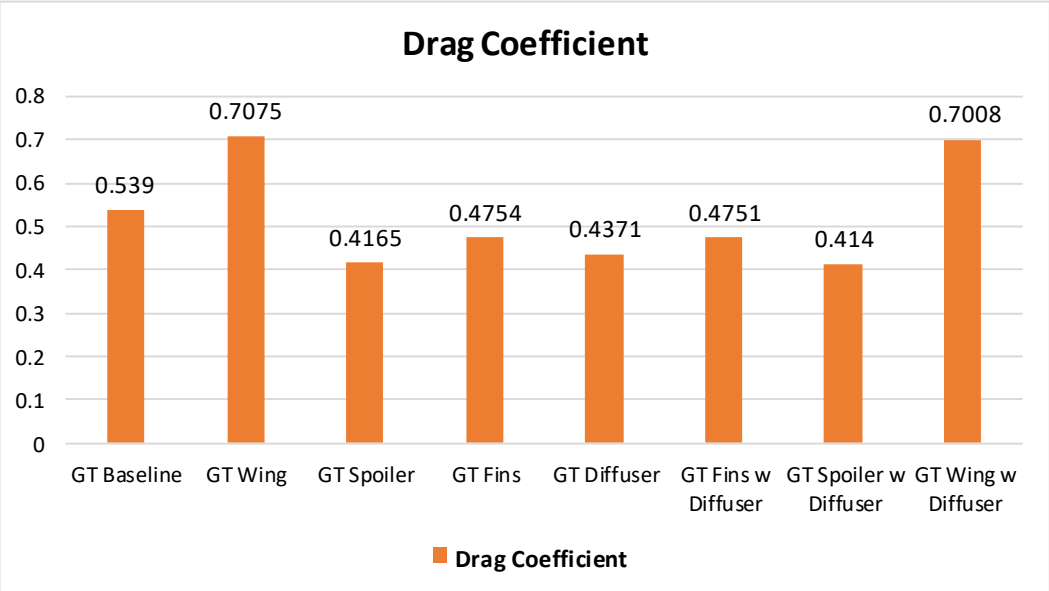


Fig. 19 Comparison of the drag coefficient for different cases

Table 2. Drag coefficient and drag force and the percentage in drag reduction from the baseline GT for different cases

S.No.	Model	Drag Coefficient	Drag Force (N)	Drag Coefficient Reduction (%)
1.	GT Baseline	0.539	1145.367	0
2.	GT Wing	0.7075	1560.828	+31.26
3.	GT Spoiler	0.4165	885.18	-22.72
4.	GT Fins	0.4754	1010.27	-11.79
5.	GT Diffuser	0.4371	928.815	-18.90
6.	GT Fins with Diffuser	0.4751	1009.656	-11.85
7.	GT Spoiler with Diffuser	0.414	879.746	-23.19
8.	GT Wing with Diffuser	0.7008	1503.464	+30.01

From Table 2 and Figures 17 & 18, it can be found that the maximum drag is in the case of the GT wing. By application of rear diffusers, the overall drag of GT Wing with the diffuser is reduced by a little margin. Secondly, the drag in the Baseline GT was further reduced by diffuser by a considerable margin. The minimum drag is observed in the case of GT Spoiler with diffuser.

Conclusions

The constant evolution in the history of vehicle aerodynamics has led to the development of certain devices which led to the enhancement of the overall aerodynamic characteristic of the vehicles. Not only it improves the efficiency of the vehicle but also reduces fuel consumption. The analysis of the baseline GT with different add on aerodynamic devices was studied by using numerical simulation in this paper. It has been found that aerodynamic drag can be influenced by using different add on devices. In consideration to reduce drag, it is favorable that the flow is attached to the vehicle's body as long as possible. A streamlined body would result in less flow separation which would cause less turbulence. In the case of GT Spoiler with Diffuser, maximum drag reduction of 23.19% is observed. Although other devices like fins also reduced drag to a much extent, they may pose a different functionality such as high-speed stability by channeling flow at rear accordingly. Wings have altogether a different function. It indeed increased the drag but its prime function is to provide downforce at the cost of increased drag, it is much like a trade-off. Diffusers on the other hand decreased the drag whenever applied in different cases. In conclusion, it may be regarded as proper optimization can lead to better aerodynamics of the vehicle in different scenarios.

Nomenclature

V	Velocity of the body
F_d	Drag force
ρ	Density of the fluid medium (air)
A	Frontal area of the body facing the fluid
C_d	Coefficient of drag of the body

Acknowledgments

The authors would like to thank the Centre for Advanced Studies and Research in Automotive Engineering (CASRAE) for providing the workspace/laboratory and the necessary equipment to carry out the required research.

Authors' contributions

The contribution of the authors to this work is equivalent. All authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

All data generated or analyzed during this study are included in this published article with appropriated citations.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Department of Mechanical Engineering, Delhi Technological University, Delhi, India.

References

[1] Wang, Jianfeng, Hao Li, Yiqun Liu, Tao Liu, and Haibo Gao. 2018. "Aerodynamic Research of a Racing Car Based on Wind Tunnel Test and Computational Fluid Dynamics." MATEC Web of Conferences 153: 1–5. https://doi.org/10.1051/matecconf/201815304011.

[2] Yang, X., Zh Cai, and Q. Ye. 2019. "Aerodynamics Analysis of Several Typical Cars." Journal of Engineering Thermophysics 28 (2): 269–75. https://doi.org/10.1134/S1810232819020085.

[3] Altinisik, Armagan, Emre Kutukceken, and Habib Umur. 2015. "Experimental and Numerical Aerodynamic Analysis of a Passenger Car: Influence of the Blockage Ratio on Drag Coefficient." Journal of Fluids Engineering, Transactions of the ASME 137 (8): 1–14. https://doi.org/10.1115/1.4030183.

[4] Sudin, Mohd Nizam, Mohd Azman Abdullah, Shamsul Anuar Shamsuddin, Faiz Redza Ramli, and Musthafah Mohd Tahir. 2014. “Review of Research on Vehicles Aerodynamic Drag Reduction Methods.” International Journal of Mechanical and Mechatronics Engineering 14 (2): 35–47.

[5] Chowdhury, Harun, Bavin Loganathan, Israt Mustary, Hazim Moria, and Firoz Alam. 2017. “Effect of Various Deflectors on Drag Reduction for Trucks.” Energy Procedia 110 (December 2016): 561–66. https://doi.org/10.1016/j.egypro.2017.03.185.

[6] Hucho W.H.: “Aerodynamics of Road Vehicles”, Forth Edition, SAE R-117, Warrendale, Pa., USA, 1997

[7] Geropp, D., & Odenthal, H. J. (2001). Drag reduction of motor vehicles by active flow control using the Coanda effect. Experiments in Fluids, 28(1), 74–85. https://doi.org/10.1007/s003480050010.

[8] Cakir, Mustafa. 2012. “Scholar Commons CFD Study on Aerodynamic Effects of a Rear Wing/ Spoiler on a Passenger Vehicle.” http://scholarcommons.scu.edu/mech_mstr.

[9] Hamut, Halil Sadettin, Rami Salah El-Emam, Murat Aydin, and Ibrahim Dincer. 2014. “Effects of Rear Spoilers on Ground Vehicle Aerodynamic Drag.” International Journal of Numerical Methods for Heat and Fluid Flow 24 (3): 627–42. https://doi.org/10.1108/HFF-03-2012-0068.

[10] https://i.stack.imgur.com/L5rdw.jpg

[11] Kajiwara, Shinji. 2017. “Passive Variable Rear-Wing Aerodynamics of an Open-Wheel Racing Car.” Automotive and Engine Technology 2 (1–4): 107–17. https://doi.org/10.1007/s41104-017-0021-9.

[12] Dey, Shamudra, and Ranabir Saha. 2018. “CFD Study on Aerodynamic Effects of NACA 2412 Airfoil as Rear Wing on a Sports Car,” no. December: 0–6.

[13] Reddy, J Jagadeep, and B Tech Mech. 2006. “Finding the Optimum Angle of Attack for the Front Wing of an F1 Car Using Cfd” 2006: 29–34.

[14] https://www.lamborghini.com

[15] https://www.koenigsegg.com

[16] Rakibul Hassan, S. M., Toukir Islam, Mohammad Ali, and Md Quamrul Islam. 2014. “Numerical Study on Aerodynamic Drag Reduction of Racing Cars.” Procedia Engineering 90: 308–13. https://doi.org/10.1016/j.proeng.2014.11.854.

[17] Hu, Xingjun, Rui Zhang, Jian Ye, Xu Yan, and Zhiming Zhao. 2011. “Influence of Different Diffuser Angle on Sedan’s Aerodynamic Characteristics.” Physics Procedia 22: 239–45. https://doi.org/10.1016/j.phpro.2011.11.038.

[18] Peddie, Kevin M, and Luis F Gonzalez. 2004. “CFD Study on the Diffuser of a Formula 3 Racecar.” Orbit The University of Sydney Undergraduate Research Journal 1 (1): 18–35.

[19] Alkan, Bugra. 2019. “Aerodynamic Analysis of Rear Diffusers for a Passenger Car by Using CFD Bir Binek Araç Için Arka Difüzör Aerodinamiğinin Hesaplamalı Akışkanlar,” no. October 2013.

AIP Conference Proceedings

Effect of concentration of alumina nanoparticles on performance and emissions of biodiesel fueled compression ignition engine

Cite as: AIP Conference Proceedings **2259**, 020013 (2020); https://doi.org/10.1063/5.0015783
Published Online: 03 September 2020

Shashank Bansal, Yash Dagar, Yash Kumar Singh, and Naushad A. Ansari

View Online Export Citation

ARTICLES YOU MAY BE INTERESTED IN

Transverse strength determination of the composite laminates
AIP Conference Proceedings **2259**, 020011 (2020); https://doi.org/10.1063/5.0016194

Effect of speed and feed rate on dry and wet turning of aluminum
AIP Conference Proceedings **2259**, 020016 (2020); https://doi.org/10.1063/5.0016144

Spatial pollution rose pattern of SO_2 in the vicinity of a typical thermal power station in India
AIP Conference Proceedings **2259**, 020017 (2020); https://doi.org/10.1063/5.0023651

Lock-in Amplifiers up to 600 MHz
Watch



AIP Publishing

© 2020 Author(s).

Effect of Concentration of Alumina Nanoparticles on Performance and Emissions of Biodiesel Fueled Compression Ignition Engine

Shashank Bansal[a)], Yash Dagar, Yash Kumar Singh, Naushad A. Ansari

Delhi Technological University, Delhi (110042), India

[a)]Corresponding author: shashankbansal98@gmail.com

Abstract. Driven by the exacerbating urbanization and population, there has been a shoot up in the global energy demand over the past few years. The technological exploration efforts in the field of biofuels have proven to be a promising alternative to the slumping fossil reserves. The main aim of this work is to inspect the potential of palm oil as a biofuel with aluminum oxide nanoparticles as an additive in it. The experimental work was carried out on a single-cylinder four-stroke CI engine running at a constant speed of 1500 rpm at varying loads using palm oil biodiesel blends and the emissions and performance parameters were compared with that of neat diesel. The results concluded a significant increase in BTE with increase in alumina concentration in the blend owing to catalytic behavior of alumina which accelerates the combustion process. Furthermore, reductions in HC and CO emissions were also recorded.

Keyword: - Diesel engine, biodiesel, emission, nanoparticles

1 INTRODUCTION

World energy demand is multiplying exponentially every day due to the proliferating urbanization, improved living standards and increasing population. For the technological, socioeconomical and industrial development of any country, energy is the major input. Most of the energy demands are met by petroleum derived fuels. Recent stats show that roughly 80% of the total energy consumed is retrieved from fossil fuels and 58% of this consumption is exploited by transportation sector. (Prabu, 2018)However, due to extensive utilization of these fuels, they are not only exhausted but also major sources of pollution in the environment. The alarming rate at which these conventional resources are depleting has prompted an urge to switch over to renewable fuels to meet the future energy demands. Alternative fuels are less harmful to environment as well as abundantly available. Apart from these, low fuel consumption is an added advantage with the use of biofuels in CI engines. (El-Seesy, Attia, & Hesham, (2018))Addition of powder form additives which are metal based to biodiesel facilitates as a catalyst in boosting the engine combustion. (Prabu, 2018)Ang F. Chen et al (Chena, Adzmia, Adamb, Othmana, l Kamaruzzamana, & Mrwana, 2018) investigated the effect on exhaust emissions , performance and characteristics of combustion of single cylinder four stroke engine by nanoparticle in diesel fuel blends. Oxides of aluminum, silica and carbon nanotubes were used as nanoparticles in his experiment. He concluded a 19.8% decrease in BSFC and 18.8% improvement in brake thermal efficiency. The paper reported better results in case of silicon oxide when compared to $Al_2O_3$and faced stability issues with the carbon nanotubes. Prabu et al (Prabu, 2018) performed experiment on a single cylinder diesel engine to analyze its efficiency and emission characteristics with alumina and cerium oxide as additives. A 12% increase in BTE was observed owing to the enhanced combustion characteristics with the use of nanoparticles. The NO, CO and HC emissions were reduced by 12%, 30% and 60% respectively. Nour et al (Nour, EL-Seesy, Abdel-Rahman, & Bady, 2018) investigated the impact adding Al_2O_3 nanoparticles in different proportions of 25, 50, 75, and 100 mg/l to a blend of 70% diesel, 10% Jojoba biodiesel, 20% ethanol at various loads on a CI engine and reported higher peak of cylinder pressure of nanoparticle biodiesel blend compared to pure diesterol blend. However, a decrease in UHC, CO, NOx, BSFC and BTE was reported on addition of nanoparticles. Ramesh et al (Ramesh, Dhananjaya Kumar, Hemanth, Namith, Parashuram, & Sharath, (2018))

International Conference on Trends in Material Science and Inventive Materials
AIP Conf. Proc. 2259, 020013-1–020013-6; https://doi.org/10.1063/5.0015783
Published by AIP Publishing. 978-0-7354-2019-9/$30.00

prepared a blend of Diesel(80%) with Poultry Litter Oil Methyl Ester(20%) and 30mg/l alumina nanoparticles to study the performance parameters of CI engine at various loads. The results show that UHC, CO, NOx and BSFC reduced, and an increase in BTE was observed on addition on Al_2O_3 nanoparticles. El-seesy et al (El-Seesy, Attia, & Hesham, (2018)) investigated by the performance that performed that by adding Al_2O_3 nanoparticles in biodiesel blend of 20% and 80% jojoba methyl ester and pure diesel respectively in the amount of 10,20,30,40, and 50 mg/l the favorable changes can be seen in emissions where it decreases as the amount of nanoparticles increases. BSFC also decreases as the amount of aluminum oxide nanoparticles increases in the biodiesel blend as per the investigation.

2 MATERIALS AND FUEL PREPARATION

2.1 Blends Preparation

Raw palm oil is first converted into the palm oil biodiesel by the process of transesterification. In this process oil is reacted with an alcohol in the presence of catalyst and product formed is ester and glycerol which is nothing but biodiesel. After the making of biodiesel, 10% palm oil biodiesel by volume is mixed in the pure diesel. To enhance properties of the blends, some amount of Al_2O_3 nanoparticle is dissolved in it. Al_2O_3 nanoparticles are formed either by grinding powder alumina particles of a nanometer level using a planetary mill or by decomposition of fresh chemically-synthesized $Al(OH)_3$ to aluminum oxide in the rapid achievement of the decomposition temperature. It has high specific surface area and ratio of length to diameter is about 20000000:1. Apart from that Al_2O_3 nanoparticles also exhibit catalytic behavior which results in rapid combustion and reducing chemical delay. Mixing of nanoparticles in the blend is done by the use of ultrasonicator. Ultrasonicator which has the power output of 240W works on the frequency of 40 kHz. The Al_2O_3 nanoparticles are dispersed in palm oil blend using ultrasonication pulsating frequency technique which helps in preventing nanoparticles agglomeration in the blend. Temperature is kept 299K for the time period of 5 hours for total dispersion of Al_2O_3 nanoparticle in the biodiesel blends. The Al_2O_3-palm oil blends samples are then taken and stored in a container under normal conditions for a time period of 4 days so as to ensure homogenous mixture and no separation.

2.2 Engine Specifications

For the performance and emissions by blends is tested on the compression ignition engine having one cylinder, four strokes. Bore of cylinder is 87.5mm and the stroke length of the piston is 110mm which is connected to a connecting rod of length 234mm. Torque of the dynamometer which is used to calculate Brake power is connected to the engine which has the arm length of 185mm. pressure of the combustion chamber is measured by the U Tube Manometer having the range of -100mm to 100mm. A gas exhaust Analyzer is also connected to the engine to measure the different emissions at different load. It measures the amount of HC, CO, present in exhaust gas in ppm. Along with that it gives the value of lambda (λ) coefficient i.e. the amount of oxygen radicals presents in exhaust gas. By the value of λ it can be depicted that whether the air fuel mixture is lean ($\lambda>1$) or rich ($\lambda<1$).

2.3 Test Performed

To check the emissions and performance characteristics of blends, the CI engine run at a constant speed of 1500 RPM keeping compression ratio also constant for variable loads. Initially engine was started on neat diesel fuel and got warm up. When the liquid water temperature stabilized, the warming up period ended. Then the performance and emissions are calculated on constant compression ratio and variable load. Dynamometer was used to vary the load on engine and was interfaced to a control panel along with Engine, which was further connected to a digital computer. The computer uses an 'Engine Analysis Software' to record various parameters like load, temperature, fuel flow rate etc. and calculating parameters like brake thermal efficiency .The engine tested the fuel

on compression ratio of 18 with the loads varied from no load to 2KG,4KG and 6KG (full load). Then similarly the test is performed for the B10(0.2g)NP and B10(0.4g)NP. Emissions on different loads are measured by the 6-gas exhaust Analyzer. The emissions and performance parameters are based on burning of per liter of diesel and biodiesel blends

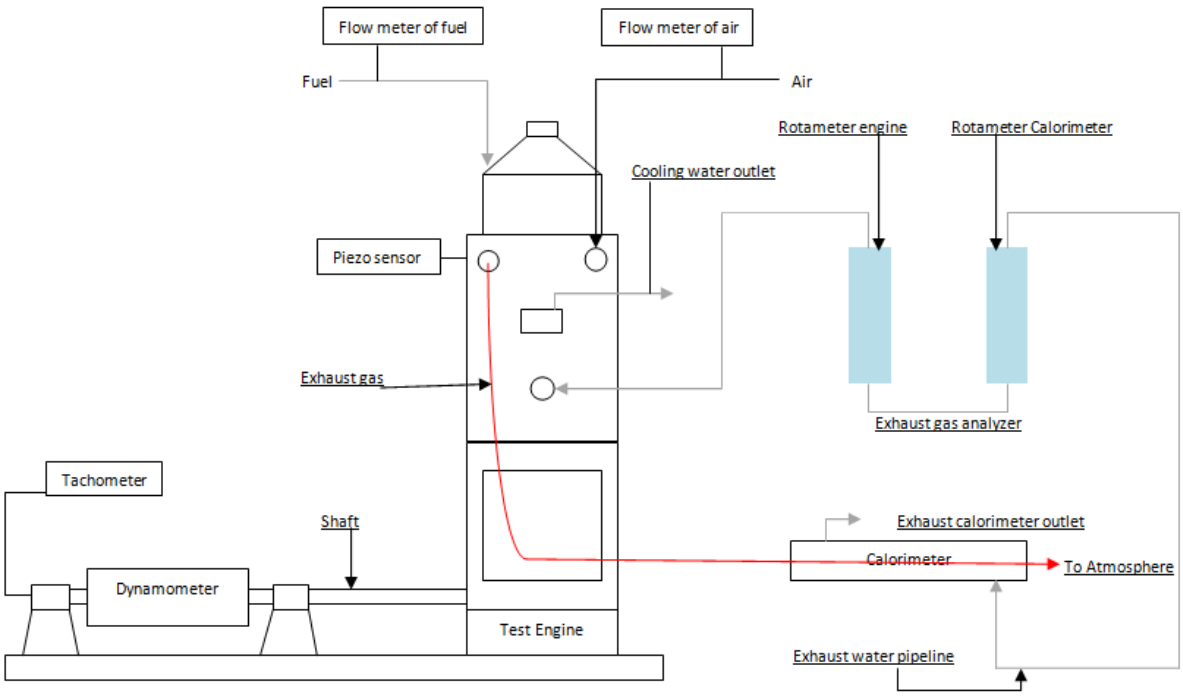


FIGURE 1. Engine Diagram

TABLE 1. Comparison of properties of diesel fuel and fuel blends

Properties	**DIESEL**	**POB10(0.2g)NP**	**POB10(0.4g)NP**
Viscosity @40°C (mm/s^2)	2.95	3.12	3.04
Cetane number	51	59	64
calorific value (MJ/Kg)	44.00	39.14	39.32
Flash point °C	76	98	120
Specific Gravity	.85	.867	.873

TABLE 2. Engine Specification

Parameters of Engine	**Specification**
Cylinder diameter	87.5 mm
Cylinder quantity	1
Strokes	4
Length of piston stroke	110 mm
Connecting rod length	234 mm
Speed of crank	1500 rpm
Rated power	3.5 kW
Arm length of dynamometer	185 mm
Orifice diameter	20 mm
Variation in injection point	0^0 to 25^0 bTDC
Range of CR	12:1 to 18:1
u- tube manometer range	10-100 mm

TABLE 3. Ultrasonicator Specifications

Ultrasonicator parameters	Specification
Power output	240 kW
Frequency	40 kHz
Temperature	299 K

3 RESULTS AND DISCUSSION

3.1 Brake Thermal Efficiency

Brake Thermal Efficiency is defined as the ratio of brake power to the heat supplied to the engine.As depicted from FIGURE 2 BTE increases with an increase in load for all blends at a constant compression ratio. Increase in BTE is observed as more amount of air is sucked in and high compression ratio results in higher efficiency.BTE is recorded to be 21.95%, 21.98% at compression ratio 18 for POB10(0.2g)NP and POB10(0.4g)NPrespectively at full load.

Effect of nanoparticles

At a constant compression ratio of 18 for POB10 blend BTE slightly increases from 0.09%, 12.62%, 17%, 21.95% at 0.2 g concentration of nanoparticles to 0.09%, 12.6%, 17.09%, 21.98% at 0.4g concentration of nanoparticles at no load, 2KG, 4KG, 6KG load respectively. The reason for increase in BTE with addition of nanoparticles may be depicted due to the catalytic behavior of Al_2O_3 nanoparticles that accelerates the combustion reaction and hence increases the efficiency. (Nour, EL-Seesy, Abdel-Rahman, & Bady, 2018)It can be seen from FIGURE 2that the slope of curve increases as the load on engine increases.

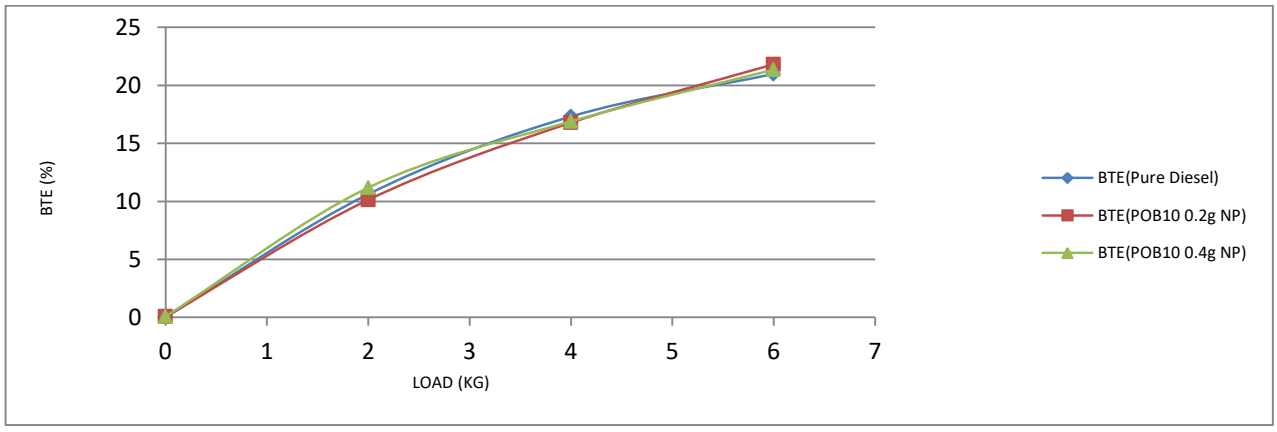


FIGURE 2. Variation of BTE vs Load

3.2 CO Emissions

Emission of carbon monoxide is due to the insufficient amount of oxygen molecule which burns in the combustion chamber and thus results in improper combustion. The intermediate chemical compound produced in the process which is nothing but CO. As depicted from FIGURE 3, at a constant compression ratio, CO emissions decreases with an increase in load for all blends. This may be due to the fact that with the increase in the amount of oxygen there is an increase in load as more amount of air is sucked in and therefore more complete combustion takes place.(Ramesh, Dhananjaya Kumar, Hemanth, Namith, Parashuram, & Sharath, (2018))

Effect of nanoparticles

At a constant compression ratio of 18 for POB10 blend CO emissions decreases from 0.04, 0.03, 0.02, 0.01 ppm at 0.2 g concentration of nanoparticles to 0.03, 0.03, 0.02, 0.01 ppm at 0.4g concentration of nanoparticles at no load, 2KG, 4KG, 6KG load respectively. The results are attributed to excellent combustion characteristics and better atomization of Al_2O_3 nanoparticles blended fuels. Addition of nanoparticles result in increase in surface area to volume ratio and improved carbon oxidation rate which in turn results in formation of CO2 instead of CO. (Ramesh, Dhananjaya Kumar, Hemanth, Namith, Parashuram, & Sharath, (2018))It can be seen from FIGURE 3that the slope of curve decreases as the load on engine increases.

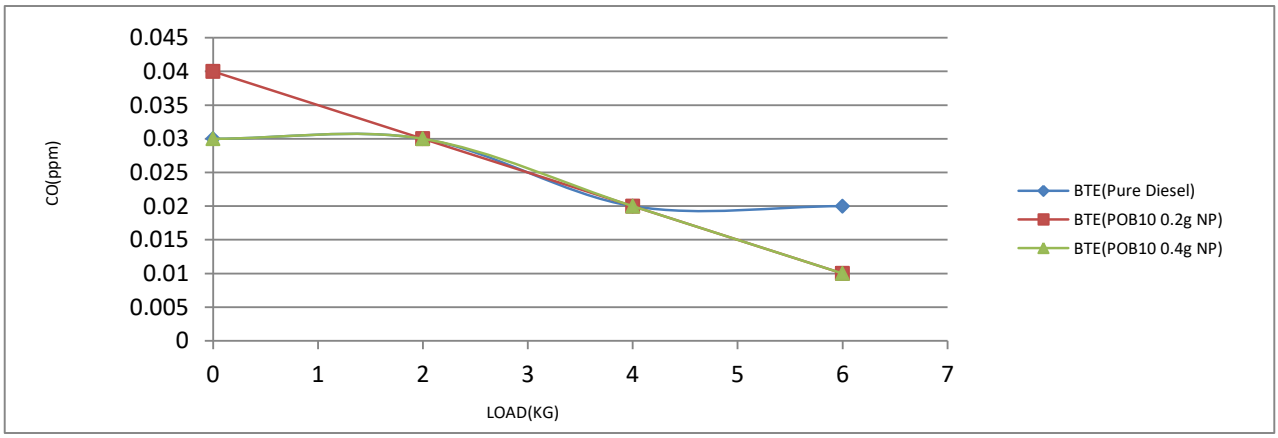


FIGURE 3. Variation of CO vs Load

3.3 HC Emissions

HC emissions are just raw fuel which does not burn properly and gets emitted from engine exhaust. There are some reasons responsible for this like low temperature and rich air fuel mixture. As depicted from the FIGURE 4 HC emission increases when load for all blends increases at a constant compression ratio. It is because of the less amount of oxygen due to the rich fuel air ratio in operation of engine. At higher engine loads, lower HC emission is due to high oxygen percentage.

Effect of nanoparticles

Drastic decrease in UBHC emissions has showed up when there is addition of alumina nanoparticle to Palm oil blends when it is compared to the pure diesel. At a constant compression ratio of 18 for POB10 blend HC emissions increases from 6, 15, 18, 21ppm at 0.2 g concentration of nanoparticles to 16, 18, 19, 21 ppm at 0.4g concentration of nanoparticles at no load, 2KG, 4KG, 6KG load respectively. The reason for that was catalytic behavior of alumina oxide. Also it could be observed the characteristics of improved ignition of alumina naoparticleS and its ignition delay shortening (El-Seesy, Attia, & Hesham, (2018)). It can be clearly interpreted from Figure 4 that 0.2g concentration of the alumina nanoparticles in the biodiesel results in lower HC emissions when compared to 0.4g concentration of the additive.

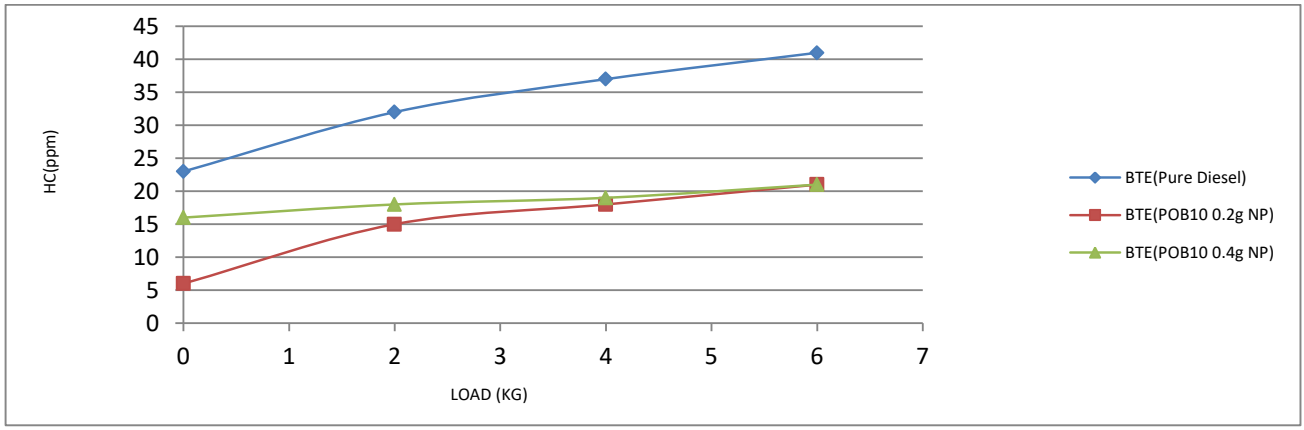


FIGURE 4. Variation of HC vs Load

4 CONCLUSION

The results from the experiment on the palm oil biodiesel blend with Al_2O_3as an additive in it are favorable for compression ratio 18, as there is increase in load, brake thermal efficiency increases from 0 to 21.95% and 21.98% for POB10(0.2g)NP and POB10(0.4g)NP respectively. There is small increase in the BTE of 0.03% if concentration of Al_2O_3 is increased in biodiesel blend.CO emission decreases up to 0.03ppm and 0.02ppm for POB10(0.2g)NP and POB10(0.4g)NP respectively when load increases at a compression ratio 18. There is slight decrease in CO emission when concentration of Al_2O_3 is increased at no load and no change at full load condition. HC emission increases with an increase in load from 6 to 21 and 16 to 21 for POB10(0.2g)NP and POB10(0.4g)NP respectively at compression ratio of 18. UBHC emissions are 6 and 16 at no load condition and no change at full load condition for POB10(0.2g)NP and POB10(0.4g)NP respectively.

REFERENCES

1. Chena, A. F., Adzmia, M. A., Adamb, A., Othmana, M. i., l Kamaruzzamana, M. K., & Mrwana, A. G. (2018). Combustion characteristics, engine performances and emissions of a diesel engine using nanoparticle-diesel fuel blends with aluminium oxide, carbon nanotubes and silicon oxide. *Energy Conversion and Management* , 461–477.
2. Effects of biodiesel on a DI diesel engine performance, emission and combustion characteristics. (3099-3105). *Fuel* , 2010.
3. El-Seesy, A. I., Attia, A. M., & Hesham, M. E.-B. ((2018)). The effect of Aluminum oxide nanoparticles addition with Jojoba methyl ester-diesel fuel blend on a diesel engine performance, combustion and emission characteristics. *Fuel* , 147–166.
4. Krupakaran, R. L., Hariprasasd, T., Gopalakrishna, A., & Babu, P. (2016). The performance and exhaust emissions investigation of a diesel engine using γ-Al2O3 nanoparticle additives to biodiesel. *Carbon Management* , 1758-3012.
5. Nour, M., EL-Seesy, A. I., Abdel-Rahman, A. K., & Bady, M. (2018). Influence of Adding Aluminum Oxide Nanoparticles to Diesterol Blends on the. *Experimental Thermal and Fluid Science* .
6. Prabu, A. (2018). Nanoparticles as additive in biodiesel on the working characteristics of a DI diesel engine. *Ain Shams Engineering Journal* , 2343-2349.
7. Ramesh, D. k., Dhananjaya Kumar, J. L., Hemanth, S. G., Namith, V., Parashuram, B. J., & Sharath, S. ((2018)). Study on effects of Alumina nanoparticles as additive with Poultry litter biodiesel on Performance, Combustion and Emission characteristic of Diesel engine. *Materials Today* , 1114–1120.

Physics of Plasmas

Electron plasma wave excitation by a q-Gaussian laser beam and subsequent electron acceleration

Cite as: Phys. Plasmas **27**, 093106 (2020); https://doi.org/10.1063/5.0007998
Submitted: 18 March 2020 . Accepted: 20 August 2020 . Published Online: 18 September 2020

Monika Yadav, Devki Nandan Gupta, and Suresh C. Sharma

View Online Export Citation CrossMark

ARTICLES YOU MAY BE INTERESTED IN

Photoelectron distribution on sunlit surface of the Moon: A formalism
Physics of Plasmas **27**, 082906 (2020); https://doi.org/10.1063/5.0016411

Direct electron bunch acceleration by Laguerre–Gauss laser pulse
Physics of Plasmas **27**, 093102 (2020); https://doi.org/10.1063/5.0015456

Magnetic reconnection and kinetic waves generated in the Earth's quasi-parallel bow shock
Physics of Plasmas **27**, 092901 (2020); https://doi.org/10.1063/5.0012443



AIP Publishing

© 2020 Author(s).

Physics of Plasmas ARTICLE scitation.org/journal/php

Electron plasma wave excitation by a q-Gaussian laser beam and subsequent electron acceleration

Cite as: Phys. Plasmas **27**, 093106 (2020); doi: 10.1063/5.0007998
Submitted: 18 March 2020 · Accepted: 20 August 2020 ·
Published Online: 18 September 2020

View Online Export Citation CrossMark

Monika Yadav,[1] Devki Nandan Gupta,[2,a)] and Suresh C. Sharma[1]

AFFILIATIONS

[1]Department of Applied Physics, Delhi Technological University, Delhi 110042, India
[2]Department of Physics and Astrophysics, University of Delhi, Delhi 110007, India

[a)]Author to whom correspondence should be addressed: dngupta@physics.du.ac.in

ABSTRACT

In this paper, we theoretically study the propagation dynamics of a q-Gaussian laser beam in a plasma by considering the relativistic and ponderomotive nonlinearities. The q-Gaussian laser beam exhibits unique characteristics while interacting with the plasma. The q-Gaussian laser beam redistributes the plasma density in a different way, which affects the laser self-focusing. A comparative study of the self-focusing for Gaussian and q-Gaussian laser beams is reported. The results obtained from numerical analysis reveal a stronger self-focusing of the q-Gaussian laser beam in plasmas, which is desirable to excite a large amplitude plasma wave for electron acceleration by extending the interaction length. We then extended this study to investigate the electron plasma wave excitation by the q-Gaussian laser beam. The electron plasma wave is driven more efficiently by the q-Gaussian laser beam. Our results show that the electron plasma wave field intensity enhances more than twofold for the q-Gaussian laser beam in comparison with the case of a Gaussian laser beam. The electron plasma wave excited by a q-Gaussian beam can accelerate the plasma electrons to higher energies. Numerical results are presented for the established set of laser and plasma parameters.

Published under license by AIP Publishing. https://doi.org/10.1063/5.0007998

I. INTRODUCTION

During the past few decades, laser technology development has drawn significant attention toward the unexplored area of laser–plasma interactions.[1–3] Its importance has been recognized in various areas such as electron acceleration,[4–8] laser driver fusion,[9] x-ray lasers,[10,11] high harmonic generation,[12] etc. For the realization of all these applications, the propagation of high-intensity laser pulses to a large distance without energy loss is highly desirable. The laser propagation without any energy loss is limited to a distance of approximately one Rayleigh length (without any optical guiding). For laser power $P > P_c$ (where P_c is the critical power), the laser beam realizes the self-focusing, which prevents the diffraction of the laser beam. Thus, the laser self-focusing is an important nonlinear phenomenon in plasmas. The laser self-focusing arises due to the modification of the dielectric property of plasma when an intense laser propagates through it.[13–15] The modification in the plasma dielectric constant arises either due to the relativistic mass variation or the ponderomotive force. The relativistic mass effect is related to the relativistic quiver velocity of the electrons due to the intense laser fields. In the relativistic-ponderomotive self-focusing mechanism, the ponderomotive force of the laser beam pushes the electrons out, whereas the ions remain immobile. Hence, the plasma electron density reduces, which further modifies the plasma dielectric constant to affect the laser beam focusing.[16–21]

A theoretical investigation of the self-focusing phenomenon was reported for the first time by Akhmanov *et al.*[22] It was then extended to a variety of media by Sodha *et al.*[23,24] and Kaw *et al.*[25] Initially, the studies on self-focusing[26–28] and electron acceleration[29,30] were restricted to the Gaussian distribution of the laser beams. However, various unexplored aspects of the beam propagation have already been revealed by considering the non-Gaussian profiles. Authors have reported laser self-focusing with different spatial profiles such as the super-Gaussian profile, Bessel profile, hollow Gaussian, elliptical Gaussian, etc.[31–42] These different laser beams show different characteristics in plasma. Patel *et al.*[43] and Nakatsutsumi *et al.*[44] have observed the intensity profile of the Vulcan Petawatt laser in their experiments. It was found that 20% energy was contained within 6.9 μm FWHM and 50% energy was contained within 16 μm FWHM on measuring the intensity distribution of Vulcan petawatt laser. The intensity distribution function of the Vulcan petawatt laser was given as $I(r) = I(0)(1 + r^2/qr_0^2)^{-q}$ (where the values of q and r_0 can be obtained by fitting the experimental data). Recently, Dwivedi and

Published under license by AIP Publishing

Malik[45] have studied the plasma channel formation by a q-Gaussian laser beam in the low-intensity regime, where the ponderomotive nonlinearity was considered for laser dynamics. They employed a q-Gaussian laser beam of intensity parameter $a_0 = 0.01$ with wavelength 10.6 μm to calculate the plasma density perturbation and subsequent self-focusing of the laser beam. At such a low-intensity, the relativistic effects are not important, hence, they were neglected in that study. We extended that work for a high-intensity laser case relevant to the laser wakefield acceleration of electrons, where both the relativistic and ponderomotive effects may be operative. Moreover, we also emphasize the previous study to explore the propagation characteristics of the q-Gaussian laser beam for electron plasma wave (EPW) excitation. The analytical study of the wakefield excitation has also been carried out using different shaped electromagnetic pulses in plasmas.[46,47] We seek a possibility to consider the q-Gaussian laser beam for plasma wave acceleration of electrons. A theoretical model supported by appropriate numerical analysis suggests to employ a q-Gaussian laser beam for strong self-focusing of the laser beam and subsequent large amplitude electron plasma wave excitation. We predict the acceleration of electrons to high energies by the electron plasma wave excited by a q-Gaussian laser beam. This paper has been divided into three parts. Section II is further divided into three parts. The first part includes the analytical model development to derive the governing equations of the self-focusing of a q-Gaussian laser beam. The second part comprises of the modified wave equation for electron plasma wave excitation and subsequent electron acceleration in the third part. We summarize the results in Sec. III.

II. THEORETICAL MODEL

A. Self-focusing of the q-Gaussian laser beam

A q-Gaussian laser beam, polarized in the x-direction, propagates along the z-direction in plasma. The electric field distribution of the q-Gaussian laser beam is given as $\boldsymbol{E}(r,z,t) = \hat{x}A(r,z,t)\exp[-i(\omega t - kz)]$, where $\hat{x}$ is the unit vector along the x-axis, ω is the laser frequency, k is the laser wave vector, and A is the amplitude of the laser field which is the function of space (r,z) and time (t). The initial intensity profile of the q-Gaussian laser beam at $z = 0$ is given by $A^2 = A_0^2(1 + r^2/qr_0^2)^{-q}g(t)$, where $g(t) = \exp(-t^2/t_0^2)$, t_0 is the laser pulse duration, and r_0 is the laser spot size. Here, q is the parameter, which indicates the deviation of the laser intensity distribution from the usual Gaussian profile. The smaller q value represents a strongly q-Gaussian profile. The q parameter tends to infinity gives the usual Gaussian distribution. The electric field of the laser beam is governed by the wave equation,

$$\nabla^2\mathbf{E} - \frac{\varepsilon}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2} = 0, \tag{1}$$

where $\varepsilon = 1 - (\omega_p^2/\omega^2) + \phi(|E|^2)$, $\omega_p = (4\pi ne^2/m)^{1/2}$ is the plasma frequency, n is the electron density, and e and m represent the electron charge and mass, respectively. The radial component of the laser field pushes the plasma electrons outward on the time-scale of plasma period ω_p^{-1}, which creates a radial space-charge field $\mathbf{E}_s = -\nabla\phi_s$. Thus, the electron density distribution may be written as

$$n_e = n_0 + \frac{1}{4\pi e}\nabla^2\phi_s. \tag{2}$$

The ponderomotive potential associated with the ponderomotive force is given by $\phi_p = -(mc^2/e)[(1 + a^2/2)^{1/2} - 1]$. In the quasi-steady state, $\phi_s = \phi_p$. Therefore, the electron density distribution can be rewritten as

$$n_e = n_0\left[1 + \frac{1}{m\omega_p^2}\left\{mc^2\nabla^2\left(1 + \frac{a^2}{2}\right)^{\frac{1}{2}}\right\}\right], \tag{3}$$

where $a^2 = (a_0^2/f^2)(1 + r^2/qr_0^2f^2)g(t)$, $a_0 = eA_0/m\omega_0 c$, and f is a dimensionless quantity (known as the beam width parameter of the laser beam). Here, we define the laser beam width as $r_0 f$. The initial laser beam width (spot size) can be obtained by taking $f = 1$. By estimating f, one may find the corresponding laser spot size variation. For example, one can say that the laser spot size reduces by decreasing the value of f. The purpose to introduce f is to make the mathematics simple and convenient.

We write the effective dielectric constant of the plasma at frequency ω as

$$\varepsilon = 1 - \frac{\omega_p^2}{\omega^2}\frac{(n_e/n_0)}{\gamma}, \tag{4}$$

where $\gamma = (1 + a^2/2)^{1/2}$. We expand the dielectric constant using Taylor expansion around $r = 0$ as $\varepsilon = \varepsilon_0 - \phi r^2/r_0^2$, where

$$\varepsilon_0 = 1 - \frac{\omega_p^2}{\omega^2}\left(1 + \frac{a_0^2}{f^2}\right)^{\frac{-1}{2}}\frac{a_0^2c^2}{r_0^2f^4\omega^2}\left(1 + \frac{a_0^2}{2f^2}\right), \tag{5}$$

$$\phi = \frac{c^2a_0^2}{\omega^2r_0^2f^6}\left(1 + \frac{a_0^2}{2f^2}\right)^{-1}\left[-\frac{a_0^2}{4f^2}\left(1 + \frac{a_0^2}{2f^2}\right)^{-1} + 2\left(1 + \frac{1}{q}\right)\right]. \tag{6}$$

Following Akhmanov *et al.*[16] and Sodha *et al.*,[17] we consider a solution of the wave equation as $\boldsymbol{E} = \boldsymbol{A}(r,z)\exp(-ikz)$, where $A(r,z)$ is the amplitude of the laser electric field. Substituting this solution in the wave equation, we get the parabolic equation as

$$-2ik\frac{\partial A}{\partial z} + \frac{\partial^2 A}{\partial r} + \frac{1}{r}\frac{\partial A}{\partial r} + \frac{\omega^2}{c^2}\varepsilon A = 0. \tag{7}$$

Using the paraxial and WKB approximations, we assume a solution of Eq. (7) as $A = A_0(r,z)\exp(-ikS)$, where A_0 and S are the real functions of r and z. Substituting this expression in Eq. (7) and separating the real and imaginary parts, we get

$$2\left(\frac{\partial S}{\partial z}\right) + \left(\frac{\partial S}{\partial r}\right)^2 = \frac{1}{k^2A}\left(\frac{\partial^2 A}{\partial r^2} + \frac{\partial A}{r\partial r}\right) - r^2\frac{\phi}{\varepsilon_0}, \tag{8}$$

$$\frac{\partial A^2}{\partial z} + \frac{\partial A^2}{\partial r}\frac{\partial S}{\partial r} + A^2\left(\frac{\partial S}{r\partial r} + \frac{\partial^2 A}{\partial r^2}\right) = 0. \tag{9}$$

To solve Eqs. (8) and (9), we assume two solutions as $S = (r^2/2)\beta(z) + \phi(z)$ and $A = (A_0/f)(1 + r^2/qr_0^2f^2)^{-q/2}$, where $\beta(z) = (1/f)df/dz$. Substituting these solutions in Eq. (9) and equating the coefficients of r^2, we obtain the expression for f as

Published under license by AIP Publishing

$$\varepsilon_0 \frac{d^2 f}{d\xi^2} = \frac{(q+4)}{q} \frac{1}{\rho^4 f^3} - \frac{\phi}{\rho^2} f, \tag{10}$$

where $\xi = z\omega/c$ is the normalized distance and $\rho = r_0\omega/c$ is the normalized laser spot size.

In Eq. (10), the first term on the right-hand side is due to the diffraction effects and the second term is the nonlinear term that is responsible for the laser self-focusing. In a general case, we use the initial boundary conditions: $f = 1$ and $df/dz = 0$ at $z = 0$ (for the initial plane wave-front) to solve this equation. For perfect guiding (where the laser spot size becomes constant), the two terms in the beam width parameter equation (representing diffraction and self-focusing effects) balance each other exactly with $f = 1$ giving a matched solution for the laser spot size as $\rho^2 = (1 + 4/q)/\phi$. However, we solve Eq. (10) for the actual evolution of the laser beam width parameter. Equation (10) is a non-linear differential equation, which governs self-focusing of a q-Gaussian laser beam when both the relativistic and ponderomotive nonlinearities are operative. We discuss the laser beam propagation characteristics using this equation. The set of laser and plasma parameters used for this calculation are as follows: $\omega = 1.8 \times 10^{15}\, s^{-1}$ (corresponding to the laser wavelength $\lambda = 1\,\mu$m), $r_0 = 9\,\mu$m, $t_0 = 40$ fs, $\omega_p/\omega_0 = 0.09$ (corresponding to the plasma electron density $n = 8 \times 10^{18}$ cm^{-3}), $v_{th}/c = 0.1$, and $a_0 = 1 - 3$ (corresponding to the laser intensity of $2.1 \times 10^{18} - 1.9 \times 10^{19}$ W/cm^2 with the laser wavelength of 1 μm). The values of r_0, v_{th}/c, and ω_p/ω are kept the same throughout the paper. In Eq. (10), the relative value of the diffractive and refractive term decides the focusing/defocusing of the laser beam. The q-value signifies the deviation of the intensity distribution of the laser beam from the Gaussian profile. For a q-Gaussian beam, the value of q is smaller. That means there is a larger deviation in pulse shape with respect to the Gaussian profile. The smaller value of q represents a q-Gaussian profile of the laser beam, whereas q tends to infinity, gives the usual Gaussian distribution. For the Gaussian profile, Eq. (10) turned out to be $\varepsilon_0 \frac{d^2 f}{d\xi^2} = \frac{1}{\rho^4 f^3} - \frac{\phi}{\rho^2} f$, where ϕ and ε_0 are modified according to the q-Gaussian profile (if $q \to \infty$). The matched laser spot size for this case would be $\rho^2 = 1/\phi$, which can also be obtained from Eq. (10). A comparison of both equations shows the evolution of the beam width parameter as the intensity distribution deviates from the usual Gaussian profile.

In Fig. 1, the variation of beam width parameter (f) with the normalized propagation distance is displayed for different values of q-parameter ($q = 0.5, 1.5, 10.5$), keeping $a_0 = 2$ and $\rho = 60$. The results show that the self-focusing effect weakens due to pulse intensity distribution as the distribution goes toward Gaussian. The sharp and faster focusing is observed for a lower q-parameter as shown in Fig. 1(a). Self-focusing length also decreases for the lower q-values. This can be explained using the fact that the different q-parameters represent different intensity distributions. For the lower q values, the non-Gaussian profile is characterized by a long and expanded tail of intensity distribution as compared to the usual Gaussian distribution. This shows the intensity shift toward the off-axial region. The diffraction of the off-axial rays is weaker comparative to the axial rays, which strengthens the plasma nonlinearity. Hence, the laser beam with the lower q-values (q-Gaussian) shows a stronger self-focusing in plasmas.

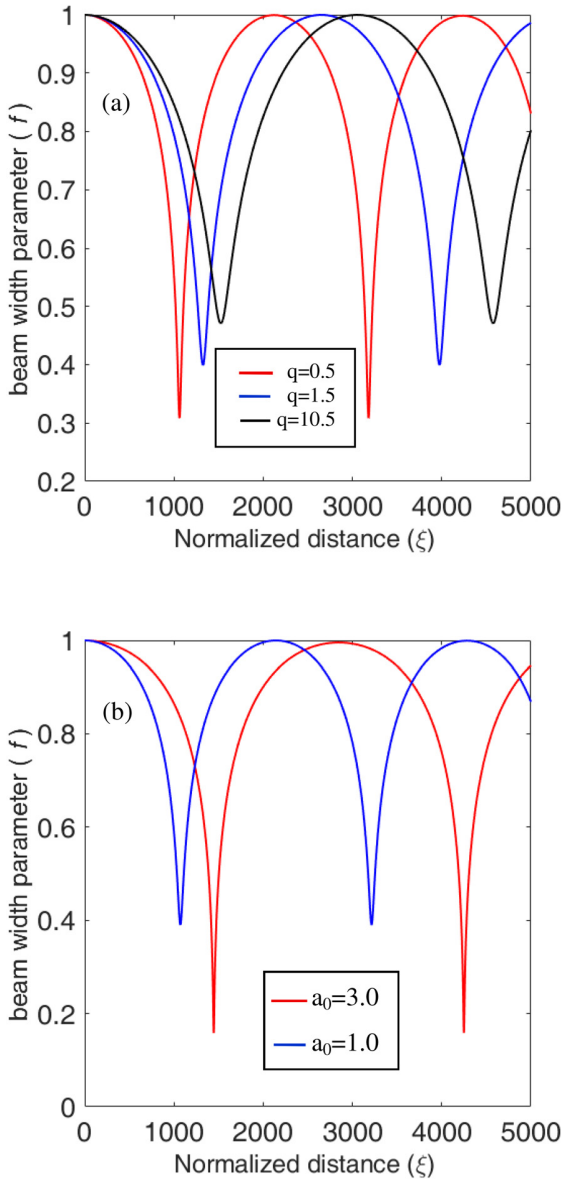


FIG. 1. The beam width parameter (f) with the normalized propagation distance (ξ) for (a) different values of q-parameter ($q = 0.5, 1.5, 10.5$) with $a_0 = 2$, $\rho = 60$, and $\omega_p/\omega_0 = 0.09$ and (b) different values of laser strength parameter ($a_0 = 1, 3$) with $q = 0.5$.

Our results show that the laser intensity also affects self-focusing of the q-Gaussian laser beam. Figure 1(b) displays the variation of the beam width parameter (f) for different a_0 ($a_0 = 1, 3$), keeping $q = 0.5$ and $\rho = 60$. The laser self-focusing enhances with the intensity of the laser beam for all q-values. The higher laser intensity reduces the diffraction term, which causes the nonlinear term to dominate. Thus, the laser self-focusing becomes stronger. However, the q-Gaussian laser beam converges

Published under license by AIP Publishing

earlier than the Gaussian beam, which makes it desirable for various applications.

B. Electron plasma wave excitation

We now employ the self-focused q-Gaussian laser pulse to excite the wakefields in plasmas. The electrostatic wakefield is resonantly excited by the laser pulse, i.e., the laser pulse length is comparable to the plasma wavelength. The ponderomotive force associated with the q-Gaussian laser pulse pushes away the plasma electrons to the low-intensity region. The ions remain immobile for such a short laser interaction so that the electrons oscillate around the initial position due to the restoring force. The density perturbation forms the electron plasma wave (EPW). In this study, the excitation of the EPW is due to the non-linear interaction of a q-Gaussian laser beam with plasmas in the presence of relativistic and ponderomotive nonlinearities. As the amplitude of the EPW depends on the background electron density, the plasma wave excitation also gets modified for the q-Gaussian laser beam. Hence, there is a strong coupling between the EPW and the q-Gaussian laser beam on account of the plasma nonlinearities. We use time averaged fluid equations and Poisson's equation to find out the electron velocity $\boldsymbol{v}$, electron density perturbation n, electric field $\mathbf{E}$, and magnetic field $\mathbf{B}$; these slowly vary over the laser period. We write the relevant equations as

$$\frac{\partial n}{\partial t} + \nabla.n\boldsymbol{v} = 0, \tag{11}$$

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v}.\nabla)\boldsymbol{v} = -\frac{e}{m}\boldsymbol{E} - \frac{e}{mc}(\boldsymbol{v}\times\boldsymbol{B}) - 2\Gamma_e\boldsymbol{v} - \frac{3K_BT_e}{m}\frac{\nabla n}{n}, \tag{12}$$

and

$$\nabla\cdot\boldsymbol{E} = -4\pi en, \tag{13}$$

where $n = n_0 + n'$ is the total plasma electron density, $v = v_0 + v'$ is the sum of the electron velocities in the electromagnetic and self-consistent field (the physical quantities denoted by the superscripts are at the plasma electron frequency), $\boldsymbol{E}$ is the sum of the electric field vectors of the electromagnetic wave and self-consistent field, Γ is the damping coefficient associated with the particle collisions (which may be approximated by $\nu_e/2$, where ν_e is the electron collision frequency in plasmas), T_e is the electron temperature, and K_B is the Boltzmann constant. The damping term is included just for general mathematical formulation of the equation of motion. We will neglect it to obtain the maximum EPW amplitude. Also, the phase velocity of the EPW is very large as compared to the thermal velocity of electrons. Hence, the damping would not be so important in this study. Using Eqs. (11)–(13), we obtain the equation for EPW at frequency $\omega = \omega_p$.

$$\frac{\partial^2 n'}{\partial t^2} + 2\Gamma\frac{\partial n'}{\partial t} - v_{th}^2\nabla^2 n' + \frac{\omega_p^2}{\gamma}\frac{n}{n_0}n' = 0, \tag{14}$$

where v_{th} is the electron thermal velocity. In this equation, the terms surviving at the electron plasma frequency are retained.

By neglecting the second and third terms in Eq. (14), one can obtain the density perturbation equation for the wakefield excitation in the bubble regime. Equation (14) is valid for a relativistic warm plasma theory. For short-pulse laser interactions, the relativistic warm plasma theory may be used such that the phase-space distribution has a small momentum spread about its mean. This allows the hierarchy of the moment equation to be treated asymptotically. No additional assumptions are necessary for closure of the fluid equations. We also apply the standard perturbation approximations to obtain the density perturbation associated with the EPW excited by the q-Gaussian laser beam. In order to get a solution of Eq. (14), we consider the quasi-static approximation, which states that the traveling wave appears static in the co-moving frame (or changes gradually in time compared to the spatial scale). Under the quasi-static approximation, the fluid quantities are assumed to be a function of the spatial variable. Moreover, by taking the quasi-static approximation, one can obtain a finite maximum density perturbation (which does not become singular). Thus, the absence of singularity confirms the validity of warm fluid equations for nonlinear wake evolution. However, the density perturbation associated with the wakefield evolution may have singularity beyond the wave breaking due to thermal effects, that we do not consider in this study. The validity of this model is for only below the wave-breaking amplitude.[48] We now assume the plane wave solution of Eq. (14) (under the WKB approximation) as

$$n' = n''(r,z)\exp\{i[\omega t - k(z + S(r,z))]\}, \tag{15}$$

where S is the eikonal of the plasma wave and ω and k are the EPW frequency and wave vector, respectively. After substituting n' in Eq. (14) and separating the real and imaginary parts, we obtain (neglecting the explicit dependence of the fluid variables on t)

$$2\frac{\partial S}{\partial z} + \left(\frac{\partial S}{\partial r}\right)^2 = \frac{1}{k^2n''}\left(\frac{\partial^2 n''}{\partial r^2} + \frac{1}{r}\frac{\partial n''}{\partial r}\right) + \frac{\omega_p^2}{k^2v_{th}^2}\left(1 - \frac{n}{\gamma n_0}\right) \tag{16}$$

and

$$\frac{\partial n''^2}{\partial z} + \frac{\partial S}{\partial z}\frac{\partial n''^2}{\partial r} + n''^2\left(\frac{\partial^2 S}{\partial r^2} + \frac{1}{r}\frac{\partial S}{\partial r}\right) + \frac{2\Gamma\omega}{kv_{th}^2}n''^2 = 0. \tag{17}$$

The variation of initial density is assumed as $n''_{z=0} = N''(1 + r^2/qh^2)^{-q/2}$, where h is the initial beam width of the EPW. The solution of Eqs. (16) and (17) can be assumed as

$$n'' = \frac{N''}{f_{pw}}\left(1 + \frac{r^2}{qf_{pw}h^2}\right)^{-q/2}\exp(-k_iz), \tag{18}$$

where f_{pw} is the beam width parameter of the EPW and $k_i = \Gamma\omega/kv_{th}^2$ is the damping factor. Substituting Eq. (18) in Eq. (16) and comparing the coefficients of r^2 on both the sides, the expression for the beam width parameter of the EPW is obtained as

$$\frac{d^2f_{pw}}{f_{pw}d\xi^2} = \frac{(q+4)}{qh^4f_{pw}^4} + \frac{c^2}{v_t^2}\left[\frac{a_0^2}{4f^8r_0^4}\left(1 + \frac{a_0^2}{2f^2}\right)^{-2} - \frac{2(q+1)a_0^2}{qf^6r_0^4}\left(1 + \frac{a_0^2}{2f^2}\right)^{-2}\right]. \tag{19}$$

Equation (18) gives the amplitude of the electron density perturbation excited by the self-focused q-Gaussian laser beam. Equation (19) shows the coupling of the EPW with the q-Gaussian laser beam under the relativistic and ponderomotive nonlinearities. Using Poisson's equation, the electrostatic field of the excited EPW can be easily assumed as

Published under license by AIP Publishing

$$\boldsymbol{E} = \hat{z}E_{pw}\exp[i(\omega t - kz)], \tag{20}$$

where

$$E_{pw} = \frac{im\omega_p^2}{ekf_{pw}}\left(1 + \frac{r^2}{qf_{pw}h^2}\right)^{-q/2}\exp(-k_i z). \tag{21}$$

Equation (21) shows the electric field amplitude of the EPW excited by the q-Gaussian laser beam. To obtain the maximum electric field, we neglect the damping term in this equation. For the laser wakefield, the radial extent of the wake is on the order of the laser spot size.[49] Thus, we solve Eq. (21) for the radial extent from 0 to r_0. Equation (21) represents the nonlinear EPW with a finite spot size (or beam width parameter). The beam width parameter of the EPW is sensitive with the spot size evolution of the driver q-Gaussian laser beam.

The interaction of a q-Gaussian laser beam with plasma under the relativistic and ponderomotive nonlinearities leads to a change in the background plasma density, which excites the EPW. The excitation of EPW is governed by Eq. (14), where the last term signifies the coupling of modified background density required for EPW excitation (when both the relativistic and ponderomotive nonlinearities are operative). Equation (19) represents the variation of the beam width parameter (f_{pw}) of EPW with $df_{pw}/dz = 0$ and $f_{pw} = 1$ at $z = 0$ (corresponding to the initial density perturbation associated with the initial EPW field amplitude). Equations (18) and (19) are numerically solved to estimate the effect of the q-Gaussian laser beam on EPW excitation. We displayed the results in terms of EPW field intensity variation with the propagation distance. The EPW field intensity is approximated by taking the square of the electric field amplitude (E_{pw}^2).

Figure 2(a) displays the EPW field intensity (in a.u.) with the normalized propagation distance for different values of q parameter, keeping $a_0 = 2$. As predicted, the EPW field intensity is higher for the q-Gaussian profile in comparison to the Gaussian profile. This is due to the stronger self-focusing of the q-Gaussian laser beam. The distance, at which the intensity peak [Fig. 2(a)] is observed, matches with the corresponding focusing points obtained in Fig. 1(a). The focused laser beam excites a stronger large amplitude EPW. To see the effect of laser intensity on the EPW excitation, we depict the EPW field intensity for different a_0, keeping $q = 0.5$. It is clearly seen that the density perturbation amplitude strongly depends on the focusing and defocusing of the laser beam. As seen from the focusing pattern of the laser beam shown in Fig. 1(b), the laser beam with higher a_0 is more focused. Thus, it is revealed that a stronger EPW may be excited if an intense laser beam is employed for the plasma interactions.

As we know, the laser beam will be Gaussian for a larger q-parameter. To see the comparison of the EPW field intensity for the Gaussian and q-Gaussian laser beam, we show the EPW field intensity for different q-parameters. Figure 3 shows the EPW field intensity for $q = 1$ (q-Gaussian laser beam) and $q = 10.5$ (approaches an ordinary Gaussian laser beam) for different a_0. We kept a_0 same for both the Gaussian and q-Gaussian laser beam for comparison. We find that the q-Gaussian laser beam excites a more intense EPW. In both the cases (Gaussian and q-Gaussian), the field intensity of EPW is enhanced with a_0. This is quite obvious as the q-Gaussian laser exhibits more self-focusing during propagation as mentioned previously. The self-focused laser beam exerts a stronger ponderomotive force on electrons. Consequently, the excited EPW shows a larger amplitude for the q-Gaussian laser beam in comparison to the Gaussian laser beam. This effect is more pronounced for higher laser intensities.

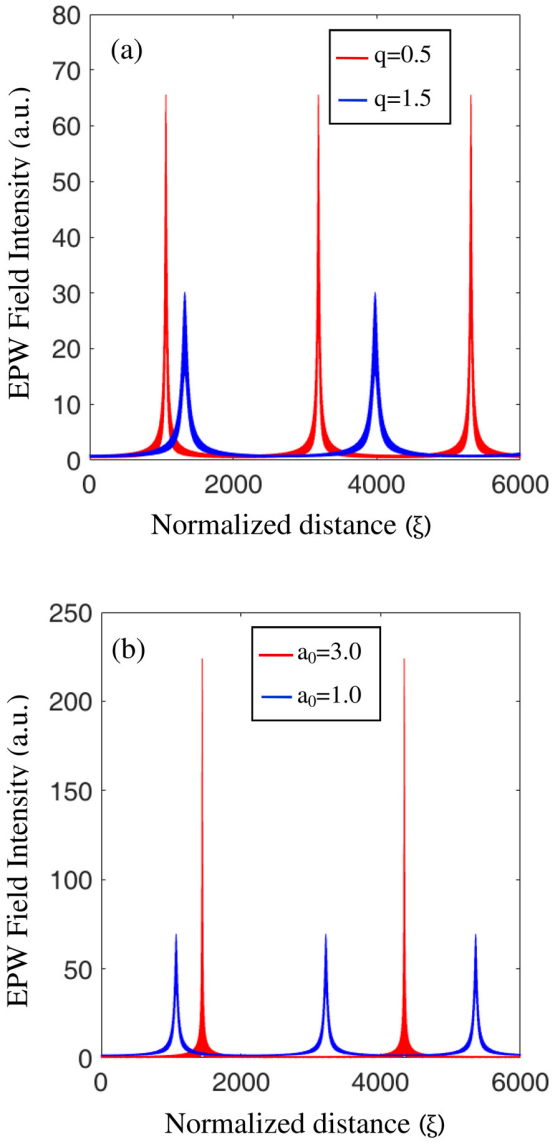


FIG. 2. The EPW field intensity (in a.u.) with the normalized propagation distance (ξ) (a) for $q = 0.5$ and 1.5, keeping the other laser parameters same as those of Figs. 1(a) and 1(b) for $a_0 = 1$ and 3, keeping the other laser parameters same as those of Fig. 1(b).

C. Electron acceleration

In Sec. II B, it was found that a large amplitude EPW can be excited by a q-Gaussian laser beam. This is resembled to the electrostatic wakefield excitation by a short-pulse laser in plasmas. The longitudinal electric field associated with the EPW can accelerate the plasma electrons. We here employ the EPW field estimated from

Published under license by AIP Publishing

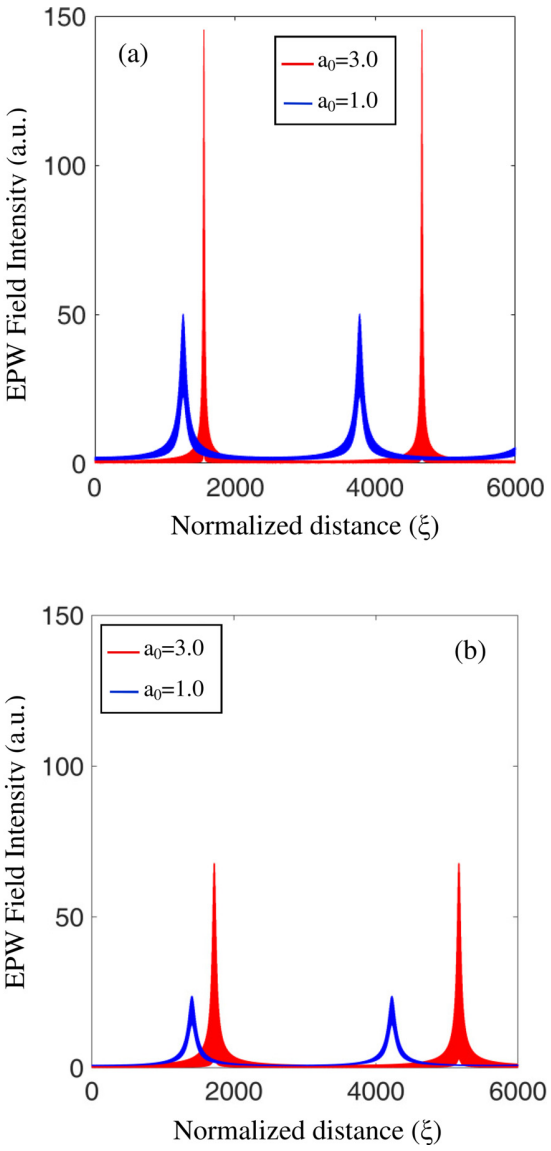


FIG. 3. The EPW field intensity (in a.u.) with the normalized distance (ξ) for (a) $q = 1.0$ and (b) $q = 10.5$ for $a_0 = 1$ and $a_0 = 3$.

Eq. (21) to obtain the electron energy gain during the acceleration. We use the equation of motion and energy equation to estimate the electron energy gain by the EPW. The equation of motion and the energy equation for the electrons are written as

$$\frac{d\boldsymbol{p}}{dt} = -e\boldsymbol{E_{pw}}, \tag{22}$$

$$\frac{d\gamma}{dt} = -\frac{e}{m_0c^2}E_{pw}v_z, \tag{23}$$

where $\boldsymbol{p} = m_0\gamma\boldsymbol{v}$ is the electron momentum and γ is the relativistic factor.

Using Eqs. (22) and (23), one can calculate the electron kinetic energy as $E_{KE} = m_0c^2\gamma$. The electric field of EPW (E_{pw}) can be calculated using Poisson's equation. We solve Eqs. (21)–(23) using appropriate numerical parameters to obtain the electron kinetic energy. It is also very clear that the electrons trapped in the accelerating phase of the large amplitude EPW are accelerated to higher energies. In Fig. 4, the maximum energy of the electrons accelerated by the EPW is shown for $a_0 = 1, 2$. The maximum electron energy reduces with the increasing q-parameter (i.e., approaching toward the Gaussian profile). That means, the electrons gain larger energy for the q-Gaussian laser beam case. The maximum electron energy gain is obtained for $q = 0.5$. The deviation of the q-parameter affects the laser self-focusing and consequently the electron energy gain. We claim that $q = 0.5$ is an optimized q-parameter to obtain the maximum electron energy gain during the EPW acceleration. For the smaller q-values, focusing has not been observed for the given set of laser and plasma parameters. Thus, the energy gain for a smaller q-parameter is not reasonable. For a smaller q-parameter, the radial variation of the laser amplitude would be large and, hence, the paraxial approximation is not valid for that propagation. The electron energy gain also depends on a_0. The maximum electron energy enhances with a_0. The reason behind this is the larger EPW amplitude excited by a stronger laser beam. We take parameter selection for an appropriate combination of parameters corresponding to the reasonable electron energy gain. The electrons gain about 45 MeV maximum energy by employing a q-Gaussian laser beam (with $q = 0.5$) of the intensity of the order of $\sim 10^{19}$ W cm^{-2}. Malik *et al.*[37] have employed various shaped laser pulses for the wakefield excitation and the electron acceleration. They predicted about 33.5 MeV electron energy gain for a rectangular–triangular laser pulse of intensity 3×10^{18} W cm^{-2}. The electron energy gain is reasonably higher for the q-Gaussian laser beam. The maximum electron energy gain may be much larger, if a higher intensity q-Gaussian laser beam is employed for the plasma interactions.

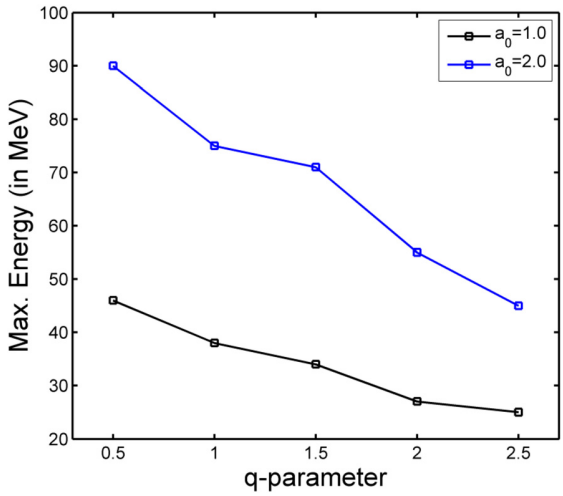


FIG. 4. The maximum electron kinetic energy (in MeV) with the q-parameter of the q-Gaussian laser beam for different laser intensity parameters $a_0 = 1$ and $a_0 = 2$ with $\omega_p/\omega_0 = 0.09$.

Published under license by AIP Publishing

As predicted, it is evident from the results that the electron energy gain depends on the self-focusing of the laser beam and the amplitude of the EPW. The electrons also gain higher energy during the acceleration for higher values of q parameter. It means that the maximum electron energy gain increases as we approach the q-Gaussian laser beam from the Gaussian laser beam. For future prospects, the laser and plasma parameters can be further optimized and new techniques can be explored to increase the maximum energy of the accelerated electrons using a q-Gaussian laser beam. Hence, the q-Gaussian laser might be of great use in the field of laser wakefield electron acceleration and might replace the Gaussian laser beam for this purpose.

III. CONCLUSION

This study examined the self-focusing phenomenon of an intense q-Gaussian laser beam and subsequently its significant effect on EPW excitation and electron acceleration. The self-focusing is found to be considerably enhanced for the lower q-values of the q-Gaussian laser beam. We observed that the self-focusing is stronger when the laser intensity is higher. We also noticed that the self-focusing has a significant influence on the EPW excitation. The EPW field intensity is seriously affected by the q-parameter of the q-Gaussian laser beam. This study shows that the field intensity of an EPW is enhanced for the lower q-values of the laser beam (i.e., for the q-Gaussian laser beam). The EPW amplitude is found to be maximum at the focal spot of the laser beam. As a result, the energy of the electrons accelerated by the EPW is also increased. The numerical results show that the electron energy gain strongly depends on the q-value of the q-Gaussian laser beam. We predict that $q = 0.5$ is the optimized q-parameter corresponding to the maximum electron energy gain for all the laser intensity parameters.

In conclusion, based on the proposed theoretical model, we suggest employing a q-Gaussian laser beam for better electron acceleration by the EPW. This study revealed the advantage of the propagation of the q-Gaussian laser beam over the Gaussian profile in plasmas under the relativistic and ponderomotive nonlinearities. This study also provides the optimized parameters of the q-Gaussian laser beam to further study the possible nonlinear processes in plasmas, which may be crucial in demonstrating the plasma channel formation by a relativistic laser beam. It might be relevant to understand the various aspects of laser–plasma interaction such as plasma heating and particle acceleration.

ACKNOWLEDGMENTS

One of the authors (M.Y.) is grateful to the University Grants Commission (UGC), Government of India for the JRF to carry out this work (No. 22/12/2013(II)EU-V). This work was financially supported by the Department of Science and Technology, Government of India under the DST-RFBR joint project (No. INT/RUS/RFBR/394).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]D. Strickland and G. Mourou, Opt. Commun. **56**, 219 (1985).
[2]J. Krall, E. Esarey, P. Sprangle, and G. Joyce, Phys. Plasmas **1**, 1738 (1994).
[3]V. K. Tripathi and C. S. Liu, Phys. Plasmas **1**, 990 (1994).
[4]A. Modena, Z. Najmudin, A. E. Dangor, C. E. Clayton, K. H. Marsh, C. Joshi, V. Malka, C. B. Darrow, C. Danson, D. Neely, and F. N. Walsh, Nature **377**, 606 (1995).
[5]G. Geddes, C. S. Toth, J. V. Tilborg, E. Esarey, C. B. Schroeder, D. Bruhwiler, C. Nieter, J. Cary, and W. P. Leemans, Nature **431**, 538 (2004).
[6]S. M. Hooker, Nat. Photonics **7**, 775 (2013).
[7]M. Yadav, S. C. Sharma, and D. N. Gupta, IEEE Trans. Plasma Sci. **46**, 2521 (2018).
[8]D. N. Gupta, K. Gopal, V. V. Kulagin, and H. Suk, Curr. Appl. Phys. **15**, 174 (2015).
[9]W. L. Kruer, *The Physics of Laser Plasma Interaction* (Addison-Wesley, New York, 1988).
[10]A. Y. Faenov, A. I. Magunov, T. A. Pikuz, I. Y. Skobelev, S. V. Gasilov, S. Stagira, F. Calegari, M. Nisoli, S. D. Silvestri, L. Poletto, P. Villoresi, and A. A. Andreev, Laser Part. Beams **25**, 267 (2007).
[11]W. Zhang and M. Y. Yu, Appl. Phys. Lett. **99**, 141501 (2011).
[12]P. Sprangle, E. Esarey, and A. Ting, Phys. Rev. Lett. **64**, 2011 (1990).
[13]H. Hora, J. Opt. Soc. Am. **65**, 882 (1975).
[14]X. L. Chen and R. N. Sudan, Phys. Rev. Lett. **70**, 2082 (1993).
[15]C. S. Liu and V. K. Tripathi, *Interaction of Electromagnetic Waves and Electron Beams with Plasmas* (World Scientific, 1994).
[16]P. Jha, N. Wadhwani, G. Raj, and A. K. Upadhyaya, Phys. Plasmas **11**, 1834 (2004).
[17]D. N. Gupta and A. K. Sharma, Phys. Scr. **66**, 262 (2002).
[18]O. Kamboj, V. Thakur, H. S. Ghotra, and N. Kant, AIP Conf. Proc. **2136**, 060011 (2019).
[19]M. A. Wani, V. Thakur, H. S. Ghotra, and N. Kant, Optik **192**, 162963 (2019).
[20]B. Hafizi, A. Ting, P. Sprangle, and R. F. Hubbard, Phys. Rev. E **62**, 4120 (2000).
[21]A. Jain, D. N. Gupta, and H. Suk, IEEE Trans. Plasma Sci. **47**, 4891 (2019).
[22]S. A. Akhmanov, A. P. Sukhorukov, and R. V. Khokhlov, Sov. Phys. JTEP **10**, 609 (1968).
[23]M. S. Sodha, A. K. Ghatak, and V. K. Tripathi, Prog. Opt. **13**, 169 (1976).
[24]M. S. Sodha, A. K. Ghatak, and V. K. Tripathi, *Self-Focusing of Laser Beams in Dielectric, Plasmas and Semiconductors* (Tata McGraw-Hill, 1974).
[25]P. Kaw, G. Schmidt, and T. Wilcox, Phys. Fluids **16**, 1522 (1973).
[26]D. N. Gupta, M. S. Hur, I. Hwang, H. Suk, and A. K. Sharma, Opt. Soc. Am. B **24**, 1155 (2007).
[27]M. S. Sodha, R. P. Sharma, and S. C. Kaushik, Plasma Phys. **18**, 879 (1976).
[28]D. N. Gupta and H. Suk, Phys. Plasmas **18**, 124501 (2011).
[29]H. S. Ghotra and N. Kant, Phys. Plasmas **23**, 013101 (2016).
[30]H. S. Ghotra and N. Kant, Opt. Commun. **383**, 169 (2017).
[31]P. Rawat, R. Gauniyal, and G. Purohit, Phys. Plasmas **21**, 062109 (2014).
[32]P. Johannisson, D. Anderson, M. Lisak, and M. Marklund, Opt. Commun. **222**, 107 (2003).
[33]B. Gaur, P. Rawat, and G. Purohit, Laser Part. Beams **34**, 621 (2016).
[34]T. S. Gill, N. S. Saini, and S. S. Kaul, Optik **115**, 493 (2004).
[35]A. Sharma and I. Kourakis, Laser Part. Beams **28**, 479 (2010).
[36]L. Wang, X. Hong, J. A. Sun, R. A. Tang, Y. Yang, W. J. Zhou, J. M. Tian, and W. S. Duan, Phys. Lett. A **381**, 2065 (2017).
[37]H. K. Malik, S. Kumar, and Y. Nishida, Opt. Commun. **280**, 417 (2007).
[38]S. Kumar, H. K. Malik, and Y. Nishida, Phys. Scr. **74**, 525 (2006).
[39]R. Kaur, T. S. Gill, and R. Mahajan, Phys. Plasmas **24**, 053105 (2017).
[40]N. Gupta and A. Singh, Optik **127**, 8542 (2016).
[41]H. K. Malik and L. Devi, Results Phys. **17**, 103070 (2020).
[42]L. Devi and H. K. Malik, Optik **207**, 164439 (2020).
[43]P. K. Patel, M. H. Key, A. J. Mackinnon, R. Berry, M. Borghesi, D. M. Chambers, H. Chen, R. Clarke, C. Damian, R. Eagleton, R. Freeman, S. Glenzer, G. Gregori, R. Heathcote, D. Hey, N. Izumi, S. Kar, J. King, A. Nikroo, A. Niles, H.-S. Park, J. Pasley, N. Patel, R. Shepherd, R. A. Snavely, D. Steinman, C. Stoeckl, M. Storm, W. Theobald, R. Town, R. Van Maren, S. C. Wilks, and B. Zhang, Plasma Phys. Controlled Fusion **47**, B833 (2005).

Published under license by AIP Publishing

[44]M. Nakatsutsumi, J. R. Davies, R. Kodama, J. S. Green, K. L. Lancaster, K. U. Akli, F. N. Beg, S. N. Chen, D. Clark, R. R. Freeman, C. D. Gregory, H. Habara, R. Heathcote, D. S. Hey, K. Highbarger, P. Jaanimagi, M. H. Key, K. Krushelnick, T. Ma, A. MacPhee, A. J. MacKinnon, H. Nakamura, R. B. Stephens, M. Storm, M. Tampo, W. Theobald, L. Van Woerkom, R. L. Weber, M. S. Wei, N. C. Woolsey, and P. A. Norreys, New J. Phys. **10**, 043046 (2008).
[45]M. Dwivedi and H. K. Malik, Opt. Lasers Eng. **129**, 106070 (2020).
[46]H. K. Malik, J. Appl. Phys. **104**, 053308 (2008).
[47]A. K. Aria and H. K. Malik, Opt. Commun. **282**, 423 (2009).
[48]E. Esarey, C. B. Schroeder, E. Cormier-Michel, B. A. Shadwick, C. G. R. Geddes, and W. P. Leemans, Phys. Plasmas **14**, 056707 (2007).
[49]E. Esarey, P. Sprangle, J. Krall, and A. Ting, IEEE Trans. Plasma Sci. **24**, 252 (1996).

Published under license by AIP Publishing

Arabian Journal for Science and Engineering
https://doi.org/10.1007/s13369-020-04841-8

RESEARCH ARTICLE-ELECTRICAL ENGINEERING

Electronically Tunable Multifunction Transadmittance-Mode Fractional-Order Filter

Garima Varshney[1] · Neeta Pandey[1] · Rajeshwari Pandey[1]

Received: 7 February 2020 / Accepted: 29 July 2020
© King Fahd University of Petroleum & Minerals 2020

Abstract
In this paper a multifunction transadmittance-mode fractional-order filter is proposed for the very first time. The proposed structure provides fractional-order low-pass filter, fractional-order high-pass filter and fractional-order all-pass filter. It uses two operational transconductance amplifier and one fractional capacitor (FC). The FC is realized using the twelfth-order continued fraction expansion-based RC ladder. Functional verification of proposed fractional-order filter is verified through PSPICE using 180-nm CMOS technology model parameters. The transfer functions sensitivities' with respect to various circuit parameters are examined through simulations. The effect of parameter variation on filter responses is studied using Monte Carlo analysis. Furthermore, the proposed filter possess electronic tunability through bias currents. Moreover, the circuit is tested experimentally using discrete components to realize the fractional-order capacitor showing great matching with the theoretical data.

Keywords Transadmittance mode · Fractional-order filter · Sensitivity · Monte Carlo Analysis

1 Introduction

Nowadays, the number of applications of fractional calculus is expanding at a great rate. This is due to the inherit capability of fractional calculus to offer more precise description of a real object than those based on classical integer methods [1–3]. Fractional-order systems are found in various branches of science, such as thermodynamics [2], electrical engineering [4], control systems [3,5–7], biomedical [8–11], mechanics, electronics and analog signal processing [4,12,13], communication and image encryption [14–16]. Therefore, immense scope of research lies for fractional-order systems at multidisciplinary level.

Filters are very important circuits in analog signal processing domain. In general, lters are categorized as current mode (CM), voltage mode (VM), transadmittance mode (TAM) and transimpedance mode (TIM) on the basis of input and output signals. In many applications voltage-mode and current-mode lters must be cascaded, which requires voltage to current converter (V–I) interface circuits to overcome these difficulties. The total efficacy of the electronic lter can be increased if signal processing can be combined with voltage to current converter interfacing. A transadmittance-mode lter is suitable for such applications and nds usage in receiver base band blocks of modern radio system [17]. Several integer-order TAM lters [18–24] are available in open literature.

The order of a filter is characterized by the slope it offers in transition band. Fractional-order filters (FOF) are more general case of classical integer-order filters, with order of the filter (n+α), where n is the integer and α is the fractional part (0< α <1). FOF offers more precise control over the transitional slope between pass band and stop band, as the slope is given by $-20(n+\alpha)$ dB/decade, while for classical filters it is -20 dB/decade. α provides an extra degree of freedom, which increase flexibility of the design [25,26]. There are two basic design approaches to develop FOFs. The first design approach is based on realizing a fractional-order element (FOE) with impedance $Z = k/s^{\alpha}$ (FO capacitor for $\alpha > 0$ and FO inductor for $\alpha < 0$). s^{α} is called as fractional Laplacian operator. This FOE structure is then replacing the respective integer-order passive element in existing circuit of integer-order filter [27–45]. The second design approach uses the integer (m)-order approximation of fractional Laplacian operator s^{α} $(0 < \alpha < 1)$, which amends the FOF of order $(1+\alpha)$ to integer-order transfer function of order $(1+m)$.

✉ Garima Varshney
garimavarshney_phd2k19@dtu.ac.in

[1] Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India

Published online: 28 August 2020

Then the final transfer function is implemented using active blocks and integrators [46–58]. A summary of FOFs based on these approaches [27–58] is given in Table 1, and following are the observations:

- The structures reported in [27–35,37–42,52–54,57] fall under category of VM, whereas those reported in [36,43–51,55,56,58] belong to CM.
- Operational amplifier (Op-Amp) [27,29,30,32,33,35,39–41,52], second-generation current conveyor (CCII) [31,34,37,38], differential difference current conveyor (DDCC) [53], operational transconductance amplifier (OTA) [36,43–45,50,56], universal voltage conveyor (UVC) [51], adjustable current amplifier (ACA) [43,50,55–57], dual output-current follower (DO-CF) [50,57], multiple output-current follower (MO-CF) [57], current feedback operational amplifier (CFOA) [54,58], current mirror [49], MOS transistors [42], differential voltage current conveyor (DVCC) [28] and current differencing buffered amplifier (CDBA) [48]-based FOFs are reported in respective references.
- The FOFs reported in [36,42–44,48–58] possess electronic tunability.

It is evident from Table 1 that there is dearth of research for TAM lters in fractional domain, and to the best of authors' knowledge, no TAM FOF is available in open literature.

This work presents a multifunction TAM FOF for the very first time. The proposed structure uses OTA as an active element and can generate fractional-order low-pass lter (FLPF), fractional-order high-pass lter (FHPF) and fractional-order all-pass lter (FAPF) of order α. The OTA is characterized by high impedance both at input and output terminals; therefore, it is the most suitable choice for implementing TAM signal processing applications.

The rest of the paper is prepared as follows. Section 2 gives a general idea on OTA and fractional-order elements, while Sect. 3 talks about general terms for FOF. Section 4 deals with the proposed structure, its stability and sensitivity analysis. Simulation results are discussed in Sect. 5 followed by experimental results in Sect. 6. Finally, the paper is concluded in Sect. 7.

2 Basic Circuit Elements

2.1 Operational Transconductance Amplifier (OTA)

The OTA provides an output current, the value of which depends upon the differential voltage applied at the input ports. The circuit symbol of OTA is shown in Fig. 1, and the relationship between input and output ports is defined in (1).

$$I_{o+} = g_m(V_+ - V_-) \tag{1}$$

Here g_m denotes the transconductance gain of OTA and depends upon bias current I_b as shown in (2)

$$g_m = \sqrt{\mu_n C_\mathrm{ox} \frac{W}{L} I_\mathrm{b}} \tag{2}$$

The symbols μ_n, C_ox, $\frac{W}{L}$ and I_b correspond mobility, oxide capacitance, aspect ratio of differential pair and bias current, respectively.

2.2 Fractional-Order Element (FOE)

The vital element in analog fractional-order circuit design is called fractance also regarded as constant phase element (CPE). The impedance of a fractance is the broad form of conventional impedances and is important for precise modeling of various circuits used in variety of applications. The generalized impedance of fractance is expressed as [59]

$$\begin{aligned} Z(s) &= ks^\alpha = k(\omega)^\alpha e^{j\frac{\alpha\pi}{2}} = k(\omega)^\alpha \left(\cos\left(\frac{\alpha\pi}{2}\right) + j\sin\left(\frac{\alpha\pi}{2}\right)\right) \\ &= |z|\angle\theta \end{aligned} \tag{3}$$

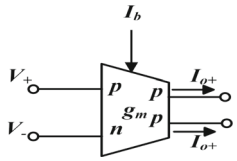
Here θ and α denote the phase and order of the fractance, respectively. θ is equal to $\alpha\pi/2$, which is constant with frequency, for a particular value of α; hence, the name is constant phase element. For $-2 < \alpha < 0$, fractance behaves as a fractional capacitor (FC), and for $0 < \alpha < 2$ fractance behaves as a fractional inductor (FI). For $\alpha = -1, 0, 1$ fractance element behaves like pure capacitor, resistor and inductor, respectively. However, for $\alpha = -2$ fractance shows quality of frequency-dependent negative resistance (FDNR) [25]. The symbol for FOC and FOI is shown in Fig. 2.

Various efforts have been made to make fractance element available commercially [60–62], but most of them are still at their developing stage. Meanwhile, various FOE approximation methods such as Carlson approximation [63], Oustaloup recursive approximation [63], Matsuda approximation [63], continued fraction expansion (CFE) [63], Chareff approximation [63], El-Khazali reduced-order approximations [64] and modified Oustaloup [65] are proposed to simulate a fractional-order element; among all the methods cited in mentioned references, the most widely used approximation method is continued fraction expansion (CFE). The approximated impedance functions obtained using FOE approximation methods can be synthesized using passive RC networks [6]/integer-order integrators [66].

Table 1 Summary of FOFs

References	Approach used	ABB used	Filter response	Order of filter	No. of elements				Electronic tunability	VM/CM/TAM	Simulation/experimental
					ABBs	R	C	C_α			
[27]	1st	Op-Amp	BP, BR	$(1+\alpha+\beta)$	2	6	1	2	No	VM	Y/Y
[28]	1st	DVCC	LP	$(\alpha+\beta)$	1	1	0	2	No	VM	Y/N
[29]	1st	Op-Amp	LP,HP,BP	$(\alpha+\beta)$	3	6	0	3	No	VM	Y/Y
[30]	1st	Op-Amp	LP, HP	$(\alpha+\beta)$	2,3	3,6	1,1	1,1	No	VM	Y/Y
[31]	1st	CCII	LP, HP, BP	$(\alpha+\beta)$	3,1	6,4	0	2,2	No	VM	Y/Y
[32]	1st	Op-Amp	LP	2α, $(\alpha+\beta)$	0,2	1,8	0	2,4	No	VM	Y/Y
[33]	1st	Op-Amp	LP	$(\alpha+\beta)$	3	6	0	2	No	VM	Y/N
[34]	1st	CCII	LP	2α	1	2	0	2	No	VM	Y/Y
[35]	1st	Op-Amp	LP, HP, BP	2α	1	5	0	2	No	VM	Y/N
[36]	1st	OTA	AP	α	2	0	0	1	Yes	CM	Y/N
[37]	1st	CCII	LP, HP, BP	α, $(\alpha+\beta)$	1,1	1,3	0	1,2	No	VM	Y/Y
[38]	1st	CCII	LP	$(\alpha+\beta)$	5	5	0	2	No	VM	Y/N
[39]	1st	Op-Amp	LP, HP, BP	$(\alpha+\beta)$	3	6	0	2	No	VM	Y/Y
[40]	1st	Op-Amp	BP	2α	4	6	2	1	No	VM	Y/Y
[41]	1st	Op-Amp	LP, BP	$(\alpha+\beta)$	3	5	1	1	No	VM	Y/Y
[42]	1st	MOS-C	LP, HP, BP	α, $(\alpha+\beta)$	2	0	0	2	Yes	VM	Y/N
[43]	1st	OTA, ACA	LP, HP, BP, BR	$(1+\alpha)$	6	0	1	1	Yes	CM	Y/N
[44]	1st	OTA	LP, BP	2α	3	0	0	2	Yes	CM	Y/N
[45]	1st	OTA	AP	α	2	0	0	1	No	CM	Y/N
[48]	2nd	CDBA	LP, HP, AP, BP, BR	$(1+\alpha)$	5	11,12	3	0	Yes	CM	Y/N
[49]	2nd	Current mirror	LP	$(1+\alpha)$	3	0	3	0	Yes	CM	Y/N
[50]	2nd	OTA, DO-CF, ACA	HP	$(1+\alpha)$	6	0	3	0	Yes	CM	Y/Y
[51]	2nd	UVC	LP, HP	$(1+\alpha)$	4, 5	10, 12	3,3	0	Yes	CM	Y/Y
[52]	2nd	Op-Amp	LP, HP, BP	$(n+\alpha)$	2	10	3	0	Yes	VM	Y/Y
[53]	2nd	DDCC	LP	$(n+\alpha)$	5	7	3	0	Yes	VM	Y/Y
[54]	2nd	CFOA	LP	$(1+\alpha)$	4	10	3	0	Yes	VM	N/Y
[55]	2nd	OTA, ACA, MO-CF	LP	$(1+\alpha)$	6	0	3	0	Yes	CM	Y/Y
[56]	2nd	OTA, ACA, CF	LP, HP	$(1+\alpha)$	7	0	3	0	Yes	CM	Y/Y
[57]	2nd	DO/MO-CF, ACA	LP	$(1+\alpha)$	3/8	3	3	0	Yes	VM	Y/Y
[58]	2nd	CFOA	LP, HP	$(n+\alpha)$	4	9,8	3,4	0	Yes	CM	Y/Y
Proposed	1st	OTA	LP, HP, AP	α	2	0	0	1	Yes	TAM	Y/Y

Fig. 1 Circuit symbol of dual output OTA



(a) OTA circuit symbol

Fig. 2 Symbol of **a** FOC, **b** FOI

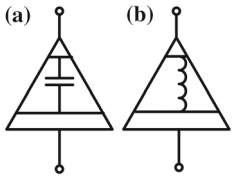


Table 2 Condition for values of a, b, c and the corresponding FOFs

Value of a, b, c	Type of FOF
$b = 0$	FLPF
$c = 0$	FHPF
$b = -ac$	FAPF

3 Fractional-Order Filters (FOF)

In this work a fractional-order TAM filter has been presented. Thus, for continuity of reading this section briefly reproduces the concept of generalized α-order filter and the related equations which have been described in detail in [25].

The generalized transfer function of α-order ($0 < \alpha < 1$) FOF is given by (4)

$$T(s)^{\alpha} = \frac{bs^{\alpha} + c}{as^{\alpha} + 1} \tag{4}$$

Here coefficients a, b and c are the constant terms, by selecting suitable values of a, b and c, fractional-order low-pass filter (FLPF), fractional-order high-pass filter (FHPF) or fractional-order all-pass filter (FAPF) may be realized, the conditions are summarized in Table 2.

The magnitude and phase response for FOFs, represented in (4), are given below:

$$|T(s)^{\alpha}| = \sqrt{\frac{b^2\omega^{2\alpha} + c^2 + 2bc\omega^{\alpha}\cos\left(\frac{\alpha\pi}{2}\right)}{\omega^{2\alpha} + a^2 + 2a\omega^{\alpha}\cos\left(\frac{\alpha\pi}{2}\right)}} \tag{5}$$

$$\angle T(s)^{\alpha} = \tan^{-1}\left[\frac{b\omega^{\alpha}\sin\left(\frac{\alpha\pi}{2}\right)}{b\omega^{\alpha}\cos\left(\frac{\alpha\pi}{2}\right) + c}\right] - \tan^{-1}\left[\frac{\omega^{\alpha}\cos\left(\frac{\alpha\pi}{2}\right) + a}{\omega^{\alpha}\sin\left(\frac{\alpha\pi}{2}\right)}\right] \tag{6}$$

Critical frequencies for FLPF, FHPF and FAPF are defined as follows [25]:

ω_m: Magnitude response is maximum or minimum
ω_{rp}: Phase response is $\pm 90°$
ω_h: Half power frequency, i.e., magnitude response is $(1/\sqrt{2})$

4 Proposed TAM FOF

The proposed TAM α-order ($0 < \alpha < 1$) multifunction FOF is shown in Fig. 3. The output current for proposed FOF is given by (7).

$$I_o = g_{m2}\left[\frac{-g_{m1}V_{\text{in1}} + s^{\alpha}C_{\alpha}V_{\text{in2}}}{g_{m2} + s^{\alpha}C_{\alpha}}\right] \tag{7}$$

For various conditions on inputs (V_{in1} and V_{in2}), different outputs are attained, and the same is mentioned in Table 3. The critical frequencies of FLPF, FHPF and FAPF are computed and given in (8)–(10), respectively.

$$\begin{aligned}
\omega_{m(\text{FLPF})} &= \omega_{0(\text{FLPF})}\left[-\cos\left(\frac{\alpha\pi}{2}\right)\right]^{\frac{1}{\alpha}} \\
\omega_{rp(\text{FLPF})} &= \omega_{0(\text{FLPF})}\left[\frac{-1}{\cos\left(\frac{\alpha\pi}{2}\right)}\right]^{\frac{1}{\alpha}} \\
\omega_{h(\text{FLPF})} &= \omega_{0(\text{FLPF})}\left[\sqrt{1 + \cos^2\left(\frac{\alpha\pi}{2}\right)} - \cos\left(\frac{\alpha\pi}{2}\right)\right]^{\frac{1}{\alpha}}
\end{aligned} \tag{8}$$

$$\begin{aligned}
\omega_{m(\text{FHPF})} &= \omega_{0(\text{FHPF})}\left[\frac{-1}{\cos\left(\frac{\alpha\pi}{2}\right)}\right]^{\frac{1}{\alpha}} \\
\omega_{rp(\text{FHPF})} &= \omega_{0(\text{FHPF})}\left[-\cos\left(\frac{\alpha\pi}{2}\right)\right]^{\frac{1}{\alpha}} \\
\omega_{h(\text{FHPF})} &= \omega_{0(\text{FHPF})}\left[\sqrt{1 + \cos^2\left(\frac{\alpha\pi}{2}\right)} + \cos\left(\frac{\alpha\pi}{2}\right)\right]^{\frac{1}{\alpha}}
\end{aligned} \tag{9}$$

$$\begin{aligned}
\omega_{m(\text{FAPF})} &= \omega_{rp(\text{FAPF})} = \omega_{0(\text{FAPF})} \\
\omega_{h(\text{FAPF})} &= \omega_{0(\text{FAPF})}\left[2\cos\left(\frac{\alpha\pi}{2}\right) + \sqrt{4\cos^2\left(\frac{\alpha\pi}{2}\right) - 1}\right]^{\frac{1}{\alpha}}
\end{aligned} \tag{10}$$

Here $\omega_{0(\text{FAPF})} = \omega_{0(\text{FLPF})} = \omega_{0(\text{FHPF})} = \omega_0 = (g_{m2}/C_{\alpha})^{1/\alpha}$

The magnitude of FAPF at $\omega = \omega_{rp}$ has a minima if $\alpha < 1$, maxima if $\alpha > 1$ and flat if $\alpha = 1$ [25].

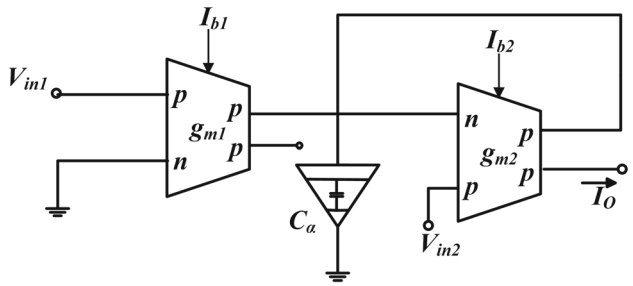


Fig. 3 OTA-based transadmittance-mode α-order FOF

4.1 Stability Analysis

Stability condition of proposed α-order TAM FOF is probed for different values of α. The stability of FOF is determined by the pole location of its characteristic equation. The procedure to analyze the stability is to first transform s-plane into fractional domain F-plane as $F = s^\alpha$, which transform the physical s-plane into $|\theta_F| < \alpha\pi$ in F-plane. The unstable region $|\theta_s| < \pi/2$ in s-plane transforms into $|\theta_F| < \alpha\pi/2$ in F-plane. Therefore, for $\alpha < 1$ region of instability in F-plane is smaller than that of in s-plane, whereas for $\alpha > 1$ the region of instability in F-plane is larger than s-plane shown in Fig. 4 [67].

The stability condition of (4) maps on to conditions of positive value of coefficient 'a' and order $\alpha < 2$ as for a positive value of 'a' and $\alpha = 2$ the system will lead to oscillation [67]. From the characteristic equation of the proposed FOF given by (7) the location of pole is attained as $s^\alpha = -g_{m2}/C_\alpha$. It is clear that, for $\alpha < 2$, the proposed α-order FOF is stable, since the location of pole lies within the stable region in F-plane.

4.2 Sensitivity Analysis

Transfer function of FOFs may change with the component's value variations. The sensitivity analysis shows the relative variation in output responses with respect to the circuit components used.

In this subsection, the sensitivities of proposed FOFs' output with respect to various circuit components have been derived mathematically.

The sensitivity of FLPF, FHPF and FAPF with respect to α, C_α, g_{m1} and g_{m2} is presented in (11), (12) and (13), respectively.

$$\begin{aligned}
S_\alpha^{\text{FLPF}} &= -\frac{\alpha s^\alpha \ln(s)}{(s^\alpha + g_{m2}/C_\alpha)} \\
S_{C_\alpha}^{\text{FLPF}} &= S_{g_{m2}}^{\text{FLPF}} = \frac{s^\alpha}{(s^\alpha + g_{m1}/C_\alpha)} \\
S_{g_{m1}}^{\text{FLPF}} &= 1
\end{aligned} \tag{11}$$

$$\begin{aligned}
S_\alpha^{\text{FHPF}} &= \frac{\frac{g_{m2}}{C_\alpha}\alpha \ln(s)}{(s^\alpha + g_{m2}/C_\alpha)} \\
S_{C_\alpha}^{\text{FHPF}} &= S_{g_{m2}}^{\text{FHPF}} = \frac{\frac{g_{m2}}{C_\alpha}}{(s^\alpha + g_{m2}/C_\alpha)} \\
S_{g_{m1}}^{\text{FHPF}} &= 0
\end{aligned} \tag{12}$$

$$S_\alpha^{\text{FAPF}} = \frac{\frac{2g_m}{C_\alpha}\alpha s^\alpha \ln(s)}{(s^{2\alpha} - g_{m2}^2/C_\alpha^2)}$$

Table 3 Various inputs and their corresponding outputs

Values of V_{in1} and V_{in2}	Output transfer function	Type of FOF
$V_{\text{in1}} = V_{\text{in}}$; $V_{\text{in2}} = 0$	$T(s)^\alpha_{\text{FLPF}} = -g_{m1}\left[\frac{1}{1+s^\alpha(C_\alpha/g_{m2})}\right]$	FLPF ($a = C_\alpha/g_{m2}, b = 0, c = -g_{m1}$)
$V_{\text{in2}} = V_{\text{in}}$; $V_{\text{in1}} = 0$	$T(s)^\alpha_{\text{FHPF}} = g_{m2}\left[\frac{s^\alpha(C_\alpha/g_{m2})}{1+s^\alpha(C_\alpha/g_{m2})}\right]$	FHPF ($a = C_\alpha/g_{m2}, b = C_\alpha, c = 0$)
$V_{\text{in1}} = V_{\text{in2}}$; $g_{m1} = g_{m2} = g_m$	$T(s)^\alpha_{\text{FAPF}} = g_m\left[\frac{s^\alpha(C_\alpha/g_{m2})-1}{s^\alpha(C_\alpha/g_{m2})+1}\right]$	FAPF ($a = C_\alpha/g_m, b = C_\alpha, c = -g_m$)

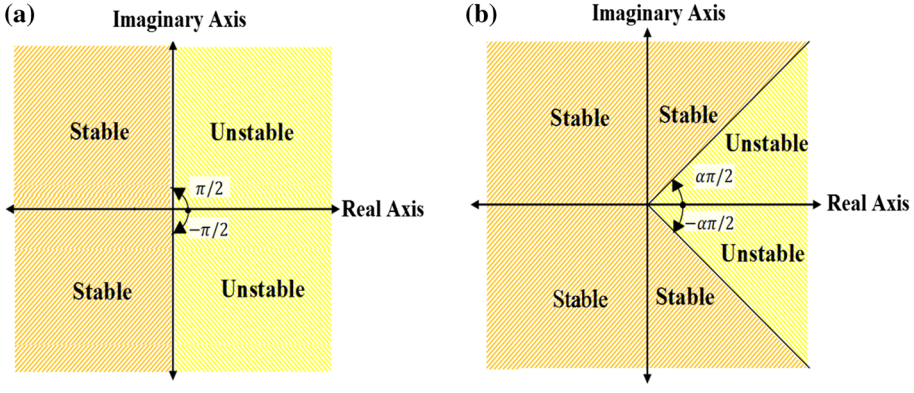


Fig. 4 Stability region in **a** s-plane, **b** F-plane

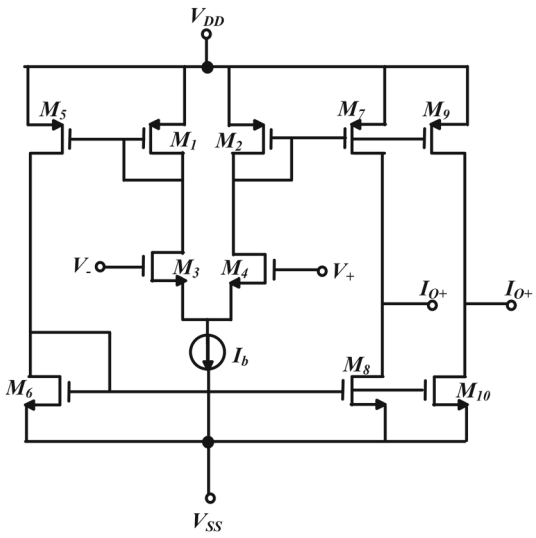


Fig. 5 CMOS structure of OTA [72]

$$S_{C_\alpha}^{\text{FAPF}} = \frac{\frac{2g_m}{C_\alpha}s^\alpha}{(s^{2\alpha} - g_{m2}^2/C_\alpha^2)}$$
$$S_{g_m}^{\text{FAPF}} = \frac{\left(s^{2\alpha} - \frac{2g_m}{C_\alpha}s^\alpha - \frac{g_m^2}{C_\alpha^2}\right)}{(s^{2\alpha} - g_{m2}^2/C_\alpha^2)} \quad (13)$$

To further demonstrate the effect of various parameters MATLAB simulations are carried out and are included in simulation section.

5 Simulation Results

Various structures of OTA are available in open literature [68–72]. To proof the concept of the proposed structure, OTA offered in [72] is used; for quick reference CMOS structure of the OTA is shown in Fig. 5. The simulations are carried out using 180-nm TSMC CMOS parameters. The supply voltages are taken as ±1.8 V. By changing the value of bias current of OTA, g_m is varied and provides electronic tunability to the circuit. The FC used in the proposed FOF is realized using RC ladder network (Foster I realization) based on 12th-order CFE approximation [73] shown in Fig. 6. The values for R_i and C_i of FC model with $\alpha = 0.5$ and $C_\alpha = 3.75\mu\mho/s^\alpha$ are $R_0 = 330$ kΩ, $R_1 = 82$ kΩ, $R_2 = 33$ kΩ, $R_3 = 12$ kΩ, $R_4 = 4.7$ kΩ, $R_5 = 2$ kΩ, $R_6 = 736$ Ω, $R_7 = 270$ Ω, $R_8 = 120$ Ω, $R_9 = 47$ Ω, $R_{10} = 8.2$ Ω, $R_{11} = 18.2$ Ω, $C_0 = 4.7$ μF, $C_1 = 3.1$ μF, $C_2 = 1$ μF, $C_3 = 470$ nF, $C_4 = 168$ nF, $C_5 = 68$ nF, $C_6 = 27$ nF, $C_7 = 10$ nF, $C_8 = 4.7$ nF, $C_9 = 1$ nF, $C_{10} = 2.2$ nF.

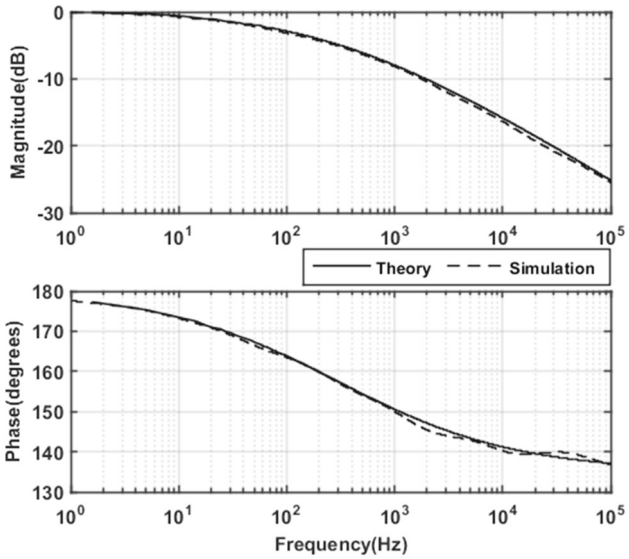


Fig. 7 Magnitude and phase response for FLPF

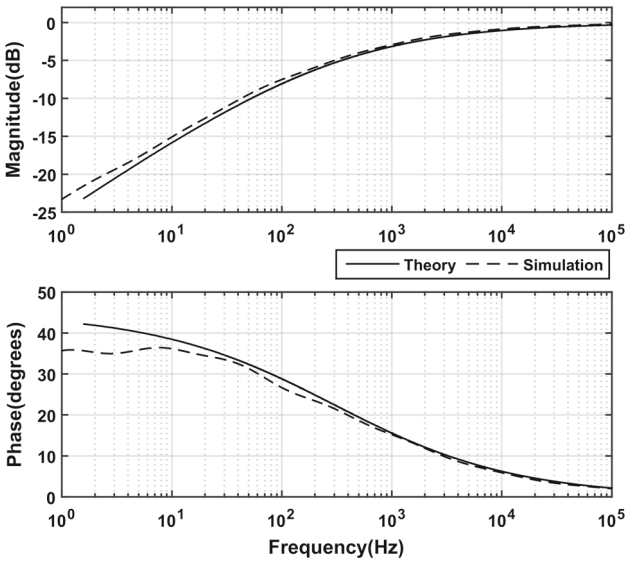


Fig. 8 Magnitude and phase response for FHPF

Simulations are carried out by fixing α at 0.5 and C_α at 3.75 $\mu\mho/s^\alpha$. I_{b1} and I_{b2} are taken equal, and the value is 15 μA; corresponding value of g_{m1} and g_{m2} is 162.3 μA/V. Normalized magnitude and phase response for FLPF, FHPF and FAPF are shown in Figs. 7, 8 and 9, respectively. The half power frequencies from the curves for FLPF and FHPF

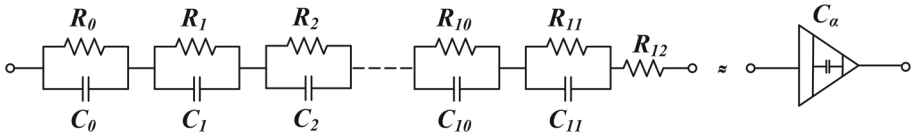


Fig. 6 Twelfth-order RC ladder network (Foster I) realizing FC

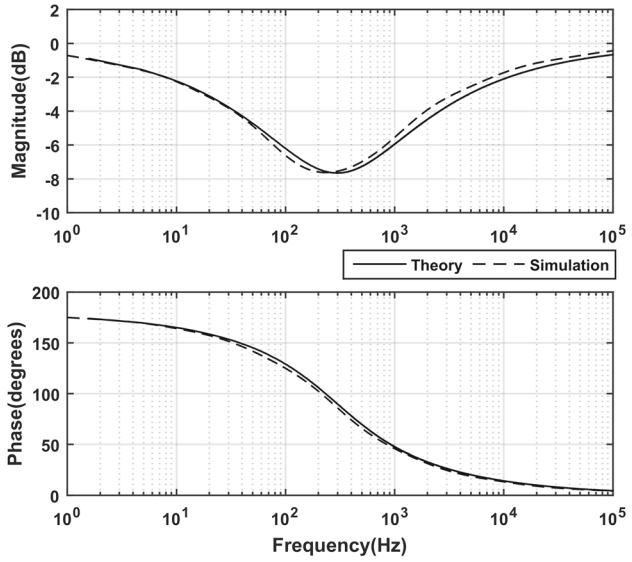


Fig. 9 Magnitude and phase response for FAPF

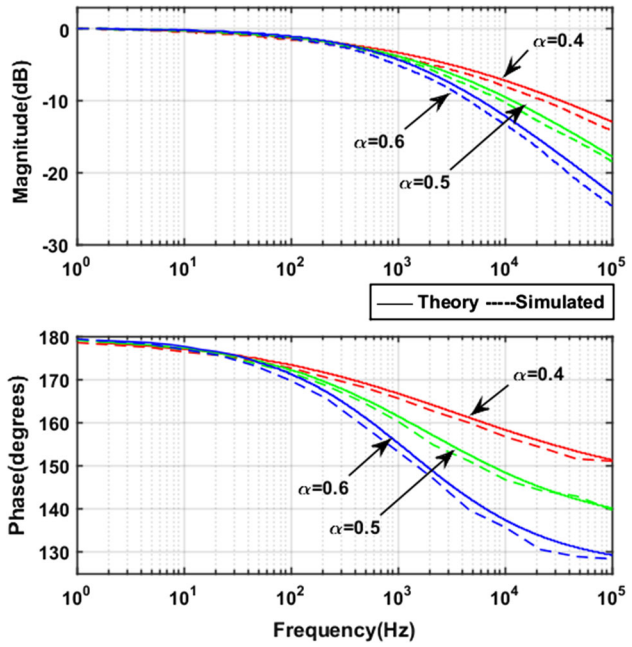
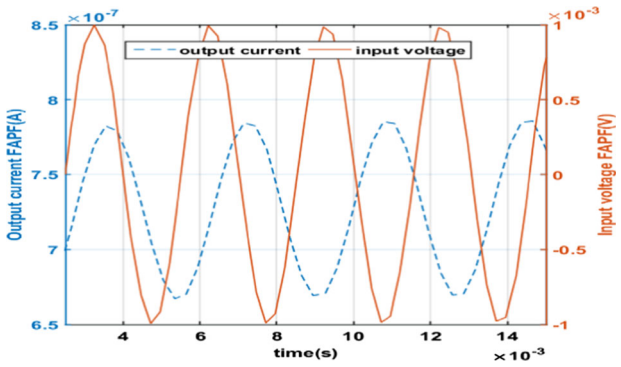
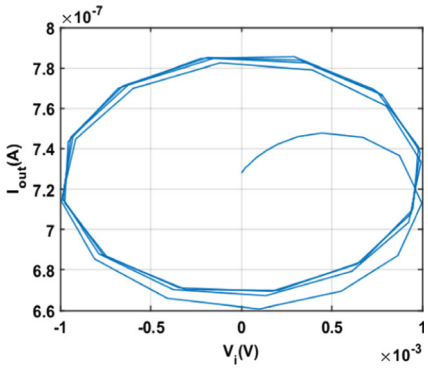


Fig. 10 Magnitude and phase response of FLPF for different values of α

are found to be 88.43 Hz and 970 Hz, respectively, while the theoretical values for the same are 83.7 Hz and 1.11 kHz. The frequency for minima for FAPF is calculated as 271.46 Hz, and the theoretical value for the same is 298.12 Hz. These observations show that both simulated and theoretical values are in close approximation. The effect of α has been studied on all the three FOFs. The magnitude and phase response of FLPF for different values of α (0.4, 0.5 and 0.6) are shown in Fig. 10, as the value of α is increased; the slope of the magnitude response is also increased as it is -20α dB/decade



(a) Transient response with 1mV, 271.46Hz input sinusoid signal



(b) Corresponding Lissajous pattern

Fig. 11 Time domain behavior of FAPF

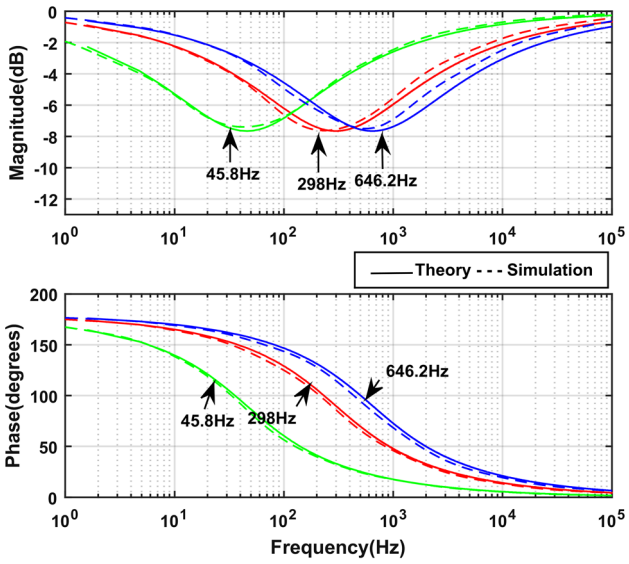


Fig. 12 Magnitude and phase response for FAPF for $I_b = 5$ μA (45.8 Hz), 15 μA (298 Hz) and 25 μA (646.2 Hz), ensuring electronic tunability

for α-order FOF. Other responses are omitted for the sake of brevity.

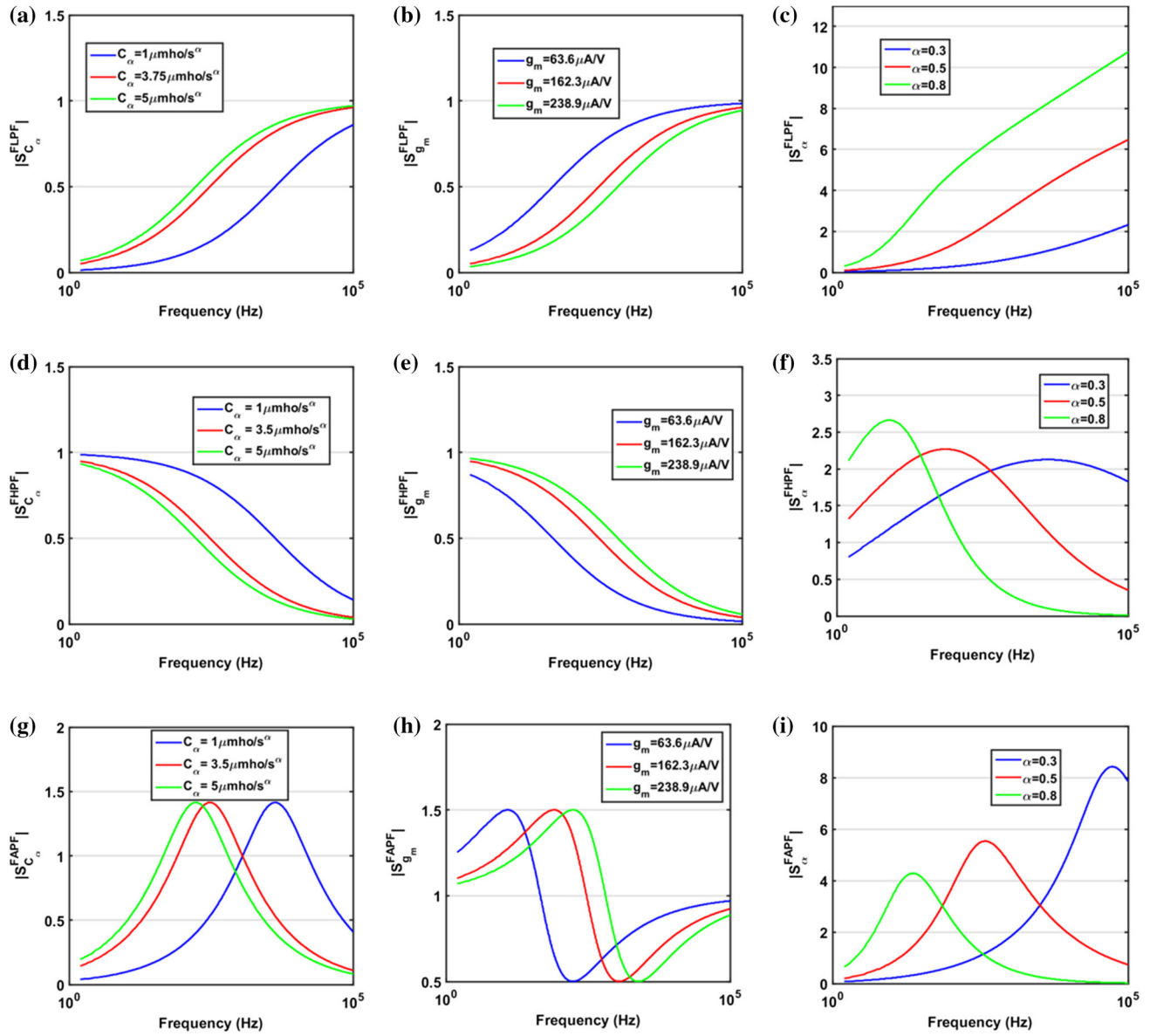
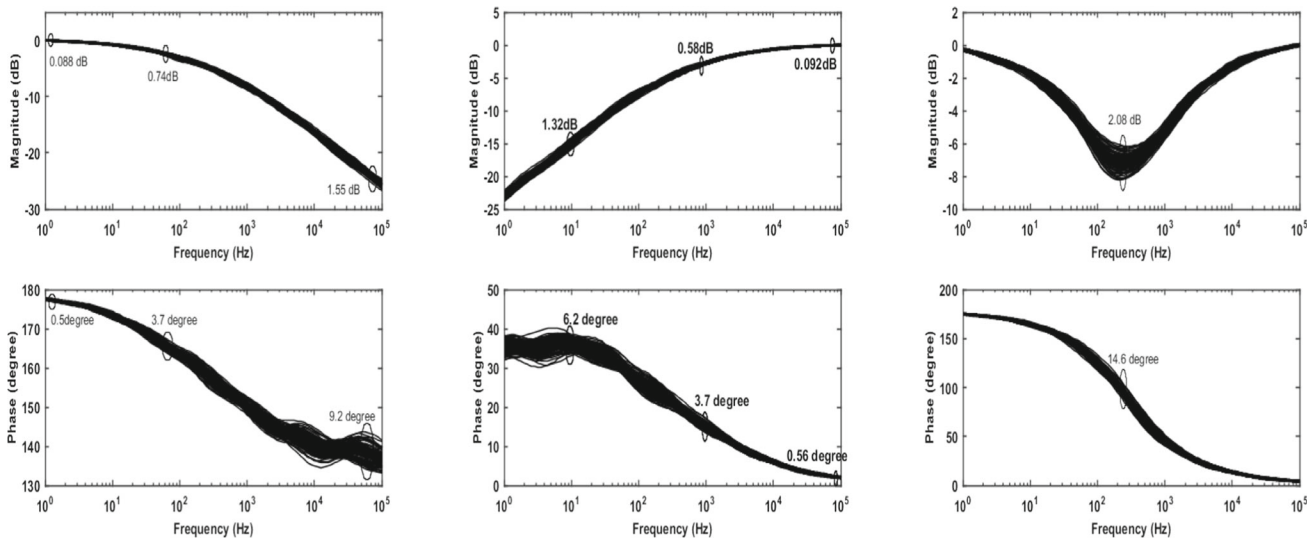


Fig. 13 Sensitivity: **a** $|S_{C_\alpha}^{\mathrm{FLPF}}|$, **b** $|S_{g_{m2}}^{\mathrm{FLPF}}|$, **c** $|S_{\alpha}^{\mathrm{FLPF}}|$, **d** $|S_{C_\alpha}^{\mathrm{FHPF}}|$, **e** $|S_{g_{m2}}^{\mathrm{FHPF}}|$, **f** $|S_{\alpha}^{\mathrm{FHPF}}|$, **g** $|S_{C_\alpha}^{\mathrm{FAPF}}|$, **h** $|S_{g_m}^{\mathrm{FAPF}}|$, **i** $|S_{\alpha}^{\mathrm{FAPF}}|$

Time domain behavior for proposed FAPF is also observed for sinusoidal input signal of 1 mV, 271.46 Hz and $\alpha = 0.5$ in Fig. 11a, and corresponding Lissajous pattern is also depicted in Fig. 11b. As the frequency of the input sinusoid is equal to the pole frequency of the FAPF which is also the right phase frequency, the Lissajous pattern shows 90° phase shift between input and output. To ensure electronic tunability, magnitude and phase response for FAPF for $I_{\mathrm{b}} = 5$ μA, 15 μA and 25 μA are plotted in Fig. 12, and the corresponding pole frequencies (f_0) are 45.8 Hz, 298 Hz and 646.2 Hz, respectively. With the change in pole frequency the phase plot would also shift in frequency axis while maintaining its shape as shown in Fig. 12.

The sensitivities of proposed FLPF, FHPF and FAPF's transfer functions with respect to frequency are studied using MATLAB tool. The sensitivity equations for FLPF, FHPF and FAPF are given in (10), (11) and (12), respectively, and plotted in Fig. 13. To plot sensitivity with respect to C_α, three different values of C_α are taken as 1 $\mu\mho/s^\alpha$, 3.75 $\mu\mho/s^\alpha$ and 5 $\mu\mho/s^\alpha$ while keeping α and g_m as constant at 0.5 and 162.3 μA/V, respectively. The corresponding simulated plot for FLPF, FHPF and FAPF is shown in Fig. 13a, d, g, respectively. Sensitivity variation curves of proposed FOFs against g_m (for g_m = 63.6 μA/V, 162.3 μA/V and 238.9 μA/V) while α is chosen as 0.5 and C_α as 3.75 $\mu\mho/s^\alpha$ are presented in Fig. 13b, e, h for FLPF, FHPF and FAPF, respectively. The variation curves for sensitivity with respect to α are plotted in



(a) FLPF **(b)** FHPF **(c)** FAPF

Fig. 14 Magnitude and phase responses under Monte Carlo analysis for proposed TAM FOF

Fig. 13c, f, i for FLPF, FHPF and FAPF, respectively, keeping C_α as 3.75 $\mu\mho/s^\alpha$ and g_m as 162.3 μA/V. The different values of α are taken as 0.3, 0.5 and 0.8.

It is witnessed that sensitivity with respect to C_α is lowest in pass band and highest in stop band for both FLPF and FHPF and within unity which is desirable feature, and same trend is followed by the sensitivity with respect to g_m for both FLPF and FHPF. For FAPF sensitivity with respect to C_α is highest at the pole frequency (slightly higher than unity), sensitivity with respect to g_m is slightly higher than unity for lower frequencies, and eventually for higher frequency, it approaches to unity. As all the filters are quite sensitive to α, the value of α must be selected very accurately. From (11) to (13) the sensitivity of all the filters with respect to α can be depicted as high due to the natural log of frequency term in the numerator.

To get further insight to the sensitivity of transfer functions with respect to C_α, uniform Gaussian distribution of the passive components used for C_α realization is considered and Monte Carlo analysis is performed by taking 100 random samples. Magnitude and phase response for proposed TAM FOF are shown in Fig. 14, due to 5% resistance and capacitance tolerances. The maximum spread for magnitude (phase) in pass band, stop band and at half power frequency is observed to be 0.088 dB (0.5°), 1.55 dB (9.2°) and 0.74 dB (3.7°) for FLPF. Similar observations for FHPF are made, and it is found that maximum spread for magnitude (phase) in pass band, stop band and at half power frequency is 0.092 dB (0.56°), 1.32 dB (6.2°) and 0.58 dB (3.7°). For FAPF the spread for magnitude (phase) is calculated at frequency of minima, and the range for the same is within 2.08 dB (14.6°). It is observed from the above analysis that the spread in magnitude and phase is within a range which verify the smaller component sensitivity of the proposed TAM FOF. The observations from the sensitivity curves and Monte Carlo analysis for FLPF, FHPF and FAPF responses with respect to C_α are summarized in Table 4. It may be noted that sensitivity has a direct bearing on spread in Monte Carlo analysis. A higher sensitivity leads to higher spread, whereas lower sensitivity results in lower spread.

6 Experimental Verification

The proposed configuration is verified experimentally also using commercially available OTA IC LM13700N with supply voltage of ± 10 V. The FC is realized using ladder structure of Fig. 6 with component values as specified in Sect. 5. The experimental setup for proposed TAM FOF is shown in Fig. 15. The transconductance gains of OTAs are set at $g_{m1} = g_{m2} = 0.46$ mA/V. The frequency response, transient response and Lissajous pattern for FLPF, FHPF and FAPF are shown in Table 5. The input signal frequencies for obtaining transient responses for FLPF and FHPF are so chosen that they fall in their respective pass bands, whereas for FAPF that frequency is selected at which minima is obtained. The input signal frequencies for FLPF and FHPF are chosen as 100 Hz and 100 kHz, respectively. For FAPF the input signal frequency is taken as 4.6 kHz which is the frequency of minima for FAPF.

It may be observed from the results shown in Table 5 that the experimental frequency responses for all the filter types are in close approximation to theoretical responses. The transient response for FLPF shows a phase difference of 177° as against the theoretical value of 174.6°. For FHPF phase

Table 4 Summary of sensitivity and Monte Carlo analysis of FOFs

Filter response	Sensitivity from sensitivity analysis		Spread region in Monte Carlo analysis	
	Maximum	Minimum	Maximum	Minimum
FLPF	Stop band	Pass band	Stop band	Pass band
FHPF	Stop band	Pass band	Stop band	Pass band
FAPF	Near pole frequency	Lower and higher frequencies	Near pole frequency	Lower and higher frequencies

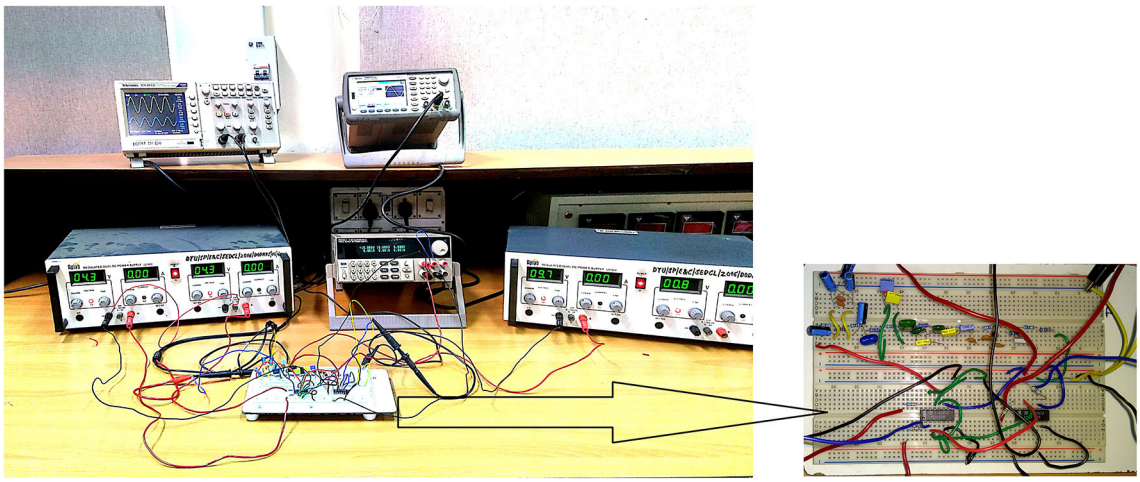
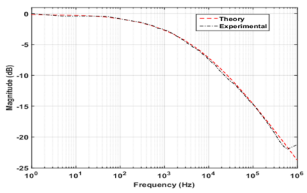
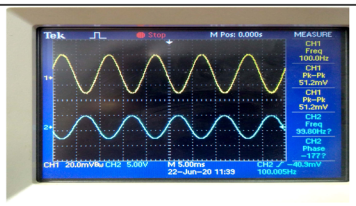
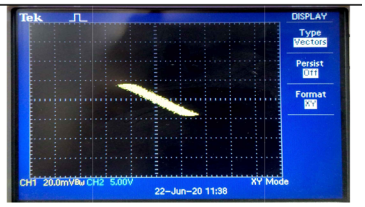
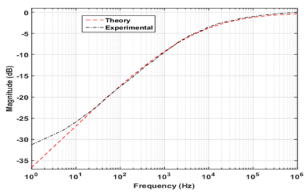
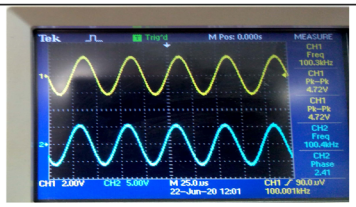
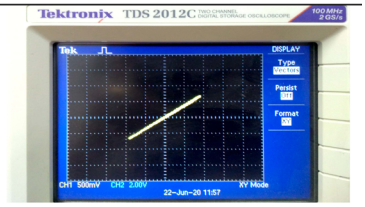
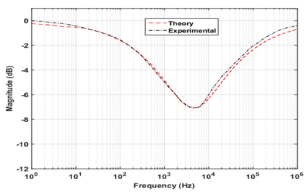
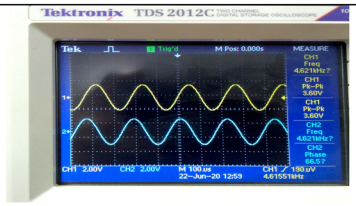
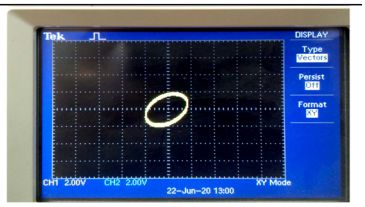


Fig. 15 Hardware setup

Table 5 Experimental results for proposed TAM FOF

Filter type	Frequency response	Transient response	Lissajous pattern
FLPF			
FHPF			
FAPF			

difference is recoded as 2.41° against the calculated value of 7.52°. For FAPF phase at minima is observed to be 66.5° in contrast to theoretical value of 90°. Minor deviations may be attributed to the components tolerances used for realizing the FC.

7 Conclusion

In this paper a new multifunction TAM FOF is proposed very first time. The proposed FOF offers three filter functions FLPF, FHPF and FAPF in transadmittance mode. Workability of the proposed structure is verified through SPICE as well as experimentally and found that results fit in the theoretical predictions very well. Electronic tunability is also verified through SPICE. Sensitivity of proposed FOF is studied using MATLAB tool, and also the effect of passive component variation on the frequency responses also is calculated using Monte Carlo analysis.

References

1. Debnath, L.: Recent applications of fractional calculus to science and engineering. Int. J. Math. Math. Sci. **54**, 3413–3442 (2003)
2. Elwakil, A.S.: Fractional-order circuits and systems: an emerging Interdisciplinary research area. IEEE Circuit Syst. Mag. **10**(4), 40–50 (2010)
3. Chen, Y.Q.; Petras, I.; Xue, D.: Fractional order control—a tutorial. In: Proceeding American Control Conference (ACC), pp. 1397–1411 (2009).
4. Das, S.; Pan, I.: Fractional order signal processing—introductory concepts and applications. In: Springer Briefs in Applied Sciences and Technology (2012).
5. Dorcak, L.; Valsa, J.; Gonzalez, E.; Terpak, J.; Petras, I.; Pivka, L.: Analogue realization of fractional-order dynamical systems. Entropy **15**(10), 4199–4214 (2013)
6. Podlubny, I.; Petras, I.; Vinagre, B.M.; Leary, P.O.; Dorcak, L.: Analogue realizations of fractional-order controllers. Nonlinear Dyn. **29**(4), 281–296 (2012)
7. Suksang, T.; Loedhammacakra, W.; Pirajnanchai, V.: Implement the fractional-order, half integrator and differentiator on the OTA based PIλDμ controller circuit. In: IEEE Conference on ECTICON (2012). https://doi.org/10.1109/ECTICON.2012.6254136
8. Geddes, L.A.; Baker, L.E.: Principles of Applied Biomedical Instrumentation, 3rd edn. Wiley, New York (1989)
9. Faria, A.C.; Veiga, J.; Lopes, A.J.; Melo, P.L.: Forced oscillation, integer and fractional-order modeling in asthma. Comput. Methods Programs Biomed. **128**, 12–26 (2016)
10. Bertsias, P.; Mohsen, M.; Said, L.A.; Elwakil, A.S.; Psychalinos, C.; Radwan, A.G.: Design and implementation of an optimized artificial human eardrum model. Circuits Syst. Signal Process. **39**, 3219–3233 (2020)
11. Tsirimokou, G.; Psychalinos, C.; Elwakil, A.: Design of CMOS Analog Integrated Fractional-Order Circuits: Applications in Medicine and Biology. Springer, Berlin (2017)
12. Sheng, H.; Chen, Y.Q.; Qiu, T.S.: Fractional Processes and Fractional-Order Signal Processing: Techniques and Applications. Springer, New York (2011)
13. Hamed, E.M.; Said, L.A.; Madian, A.H.; Radwan, A.G.: On the approximations of CFOA-based fractional-order inverse filters. Circuits Syst. Signal Process. **39**, 2–29 (2020)
14. Tlelo-Cuautle, E.; Pano-Azucena, A.D.; Guillén Fernández, O.; Silva-Juárez, A.: Analog/Digital Implementation of Fractional Order Chaotic Circuits and Applications. Springer, Berlin (2020)
15. Muñiz-Montero, C.; Sánchez-Gaspariano, L. A.; Sánchez-López, C.; González-Díaz, V. R.; Tlelo Cuautle, E.: On the electronic realizations of fractional-order phase-lead-lag compensators with OpAmps and FPAAs. In: Fractional Order Control and Synchronization of Chaotic Systems, vol. 688, pp. 131–164 (2017). ISBN: 978-3-319-50248-9
16. Sayed, W.S.; Radwan, A.G.: Generalized switched synchronization and dependent image encryption using dynamically rotating fractional-order chaotic systems. AEU Int. J. Electron. Commun. (2020). https://doi.org/10.1016/j.aeue.2020.153268
17. Toker, A.; Cicekoglu, O.; Ozcan, S.; Kuntman, H.: High output-impedance transadmittance type continuous-time multifunction filter with minimum active elements. Int. J. Electron. **88**, 1085–1091 (2001)
18. Nand, D.; Pandey, N.: Transadmittance mode first order LP/HP/AP filter and its application as an oscillator. IOP Conf. Ser. Mater. Sci. Eng. **225**, 012150 (2017)
19. Minaei, S.: A new high performance CMOS third generation current conveyor (CCIII) and its application. J. Electr. Eng. **85**, 147–153 (2003)
20. Cam, U.: A new transadmittance type first order all pass filter employing single third generation current conveyor. Analog Int. Circuits Signal Process. **43**, 97–99 (2005)
21. Herencsar N.; Lahiri A.; Koton J.; Sagbas M.; Ayten U. E.; Vrba K.: New MOS-C realization of transadmittance type all pass filter using modified CBTA. In: International Conference on Applied Electronics, pp. 1–4. IEEE (2011)
22. Herencsar N.; Koton J.; Vrba K.: Differential input buffered and transconductance amplifier based new transadmittance and voltage mode first order all pass filters. In: International Conference on Electrical and Electronics Engineering, ELECO-2009, pp. 256-259 (2009)
23. Herencsar N.; Koton J.; Vrba K.; B. Metin: Fully cascadable dual mode all pass filter based on single DBTA. In: 35th International Conference on Telecommunications and Signal Processing, pp. 374–377 (2012)
24. Sotner, R.; Herencsar, N.; Jerabek, J.; Vrba, K.; Dostal, T.; Jaikla, W.; Metin, B.: Novel first order all pass filter applications of z-copy voltage differencing current conveyor. Indian J. Pure Appl. Phys. **53**, 537–545 (2015)
25. Radwan, A.G.; Soliman, A.M.; Elwakil, A.S.: First-order filters generalized to the fractional domain. J. Circuits Syst. Comput. **17**(1), 55–66 (2008)
26. Radwan, A.G.; Elwakil, A.S.; Soliman, A.M.: On the generalization of second order filters to the fractional order domain. J. Circuits Syst. Comput. **18**(2), 361–386 (2009)
27. Adhikary, A.; Sen, S.; Biswas, K.: Practical realization of tunable fractional order parallel resonator and fractional order filters. IEEE Trans. Circuits Syst. I **63**(8), 1142–1151 (2016)
28. Mishra, S.K.; Gupta, M.; Upadhyay, D.K.: Active realization of fractional order Butterworth low pass filter using DVCC. J. King Saud Univ. Eng. Sci. **32**(2), 158–165 (2020)
29. Kubanek, D.; Freeborn, T.; Koton, J.: Fractional-order band-pass filter design using fractional-characteristic specimen functions. Microelectron. J. **86**, 77–86 (2019)
30. Hamed, E.M.; AbdelAty, A.M.; Said, L.A.; Radwan, A.G.: Effect of different approximation techniques on fractional-order KHN filter design. Circuits Syst. Signal Process. **37**, 5222–5252 (2018). https://doi.org/10.1007/s00034-018-0833-5
31. Ali, A.S.; Radwan, A.G.; Soliman, A.M.: Fractional order butterworth filter: active and passive realization. IEEE J. Emerg. Sel. Top. Circuits Syst. **3**(3), 346–354 (2013)
32. AbdelAty, A.M.; Soltan, A.; Waleed, A.; Radwan, A.G.: Fractional order Chebyshev-like low-pass filters based on integer order poles. Microelectron. J. **90**, 72–81 (2019)
33. Soltan, A.; Radwan, A.G.; Soliman, A.M.: Fractional order filter with two fractional elements of dependant orders. Microelectron. J. **43**(11), 818–827 (2012)

34. Said, L.A.; Ismail, S.M.; Radwan, A.G.; Median, A.H.; El Yazeed, M.F.A.; Soliman, A.M.: On the optimization of fractional order low pass filter. Circuits Syst. Signal Process. **35**(6), 2017–2039 (2016)
35. Bhaskar, D.R.; Kumar, M.; Kumar, P.: Fractional order inverse filters using operational amplifier. Analog Integr. Circuits Signal Process. **97**(1), 149–158 (2018)
36. Kamath, D.V.; Navya, S.; Soubhagyaseetha, N.: Fractional order OTA-C current-mode all-pass filter. In: Second International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, pp. 383–387 (2018)
37. Khalil, N.A.; Said, L.A.; Radwan, A.G.; Soliman, A.M.: Generalized two port network based fractional order filters. AEU Int. J. Electron. Commun. **104**, 128–146 (2019)
38. Soltan, A.; Radwan, A.G.; Soliman, A.M.: CCII based fractional lters of different orders. J. Adv. Res. **5**, 157–64 (2014)
39. Tripathy, M.C.; Biswas, K.; Sen, S.: A design example of a fractional order Kerwin–Huelsman–Newcomb biquad lter with two fractional capacitors of different order. Circuits Syst. Signal Process. **32**, 1523–36 (2013)
40. Ahmadi, P.; Maundy, B.; Elwakil, A.S.; Belostotski, L.: High-quality factor asymmetric-slope band-pass lters: a fractional order capacitor approach. IET Circuits Devices Syst. **6**, 187–97 (2012)
41. Freeborn, T.J.; Maundy, B.; Elwakil, A.S.: Fractional-step Tow–Thomas biquad lters. Nonlinear Theor. Appl. **3**, 357–74 (2012)
42. Bertsias, P.; Psychalinos, C.; Elwakil, A. S.; Maundy, B. J.: Simple multi-function fractional-order filter designs. In: IEEE conference on MOCAST (2019). https://doi.org/10.1109/MOCAST.2019.8741674
43. Dvorak, J.; Jerabek, J.; Polesakova, Z.; Langhammer, L.; Blazek, P.: Multifunctional electronically recongurable and tunable fractional-order lter. Elektronika IR Elektrotechnika **25**(1), 26–30 (2018)
44. Verma, R.; Pandey, N.; Pandey, R.: Electronically tunable fractional order filter. Arab. J. Sci. Eng. **42**, 3409–3422 (2017)
45. Bhat, M.V.; Bhat, S.S.S.; Kamath, D.V.: Gm-C current mode fractional all pass filter of order (0¡alpha¡1), In: IEEE Conference ICECA (2019). https://doi.org/10.1109/ICECA.2019.8822183
46. Mahata, S.; Saha, S.K.; Kar, R.; Mandal, D.: Optimal design of fractional order low pass butterworth lter with accurate magnitude response. Digit. Signal Process, **72**, 96–114 (2018)
47. Soni, A.; Sreejeth, N.; Saxena, V.; Gupta, M.: Series optimized fractional order low pass butterworth filter. Arab. J. Sci. Eng. (2019). https://doi.org/10.1007/s13369-019-04225-7
48. Kaur, G.; Ansari, A.Q.; Hashmi, M.S.: Analysis and investigation of CDBA based fractional-order filters. Analog Integr. Circuits Signal Process. (2020). https://doi.org/10.1007/s10470-020-01683-0
49. Tsirimokou, G.; Psychalinos, C.: Ultra-low voltage fractional-order circuits using current mirrors. Int. J. Circuits Theor. Appl. **44**, 109–26 (2016)
50. Jerabek, J.; Sotner, R.; Dvorak, J.; Langhammer, L.; Koton, J.: Fractional-order high-pass lter with electronically adjustable parameters. In: International Conference on Applied Electronics, pp. 111–116, (2016). https://doi.org/10.1109/AE.2016.7577253
51. Koton, J.; Kubanek, D.; Sladok, O.; Vrba, K.: Fractional-order low- and high-pass lters using UVCs. J. Circuits Syst. Comput. **26**(12), 1–23 (2017)
52. Maundy, B.; Elwakil, A.S.; Freeborn, T.J.: On the practical realization of higher order lters with fractional stepping. Signal Process. **91**(3), 484–91 (2011)
53. Khateb, F.; Kubanek, D.; Tsirimokou, G.; Psychalinos, C.: Fractional-order lter based on low-voltage DDCCs. Microelectron. J. **50**, 50–59 (2011)
54. Tsirimokou, G.; Koumousi, S.; Psychalinos, C.: Design of fractional-order lters using current feedback operational ampliers. J. Eng. Sci. Technol. Rev. **9**(4), 77–81 (2016)
55. Langhammer, L.; Dvorak, J.; Jerabek, J.; Koton, J.; Sotner, R.: Fractional-order low-pass lter with electronic tunability of its order and pole frequency. J. Electr. Eng. Slovak. **69**(1), 3–13 (2018)
56. Jerabek, J.; Sotner, R.; Dvorak, J.; Polak, J.; Kubanek, D.; Herencsar, N.; Koton, J.: Recongurable fractional-order lter with electronically controllable slope of attenuation, pole frequency and type of approximation. J. Circuits Syst. Comput. **26**(10), 1–21 (2017)
57. Dvorak, J.; Langhammer, L.; Jerabek, J.; Koton, J.; Sotner, R.; Polak, J.: Synthesis and analysis of electronically adjustable fractional-order low-pass lter. J. Circuits Syst. Comput. **27**(2), 1850032-(1-18) (2018)
58. Verma, R.; Pandey, N.; Pandey, R.: CFOA based low pass and high pass fractional step filter realizations. AEU Int. J. Electron. Commun. **99**, 161–176 (2019)
59. Nakagava, M.; Sorimachi, K.: Basic characteristics of a fractance device. IEICE Trans. Fundam. Electron. Commun. Comput. Sci. **75**(12), 1814–1818 (1992)
60. Agambayev, A.; Farhat, M.; Patole, S.P.; Hassan, A.H.; Bagci, H.; Salama, K.N.: An ultra-broadbandsingle-component fractional-order capacitor using MoS2-ferroelectric polymer composite. Appl. Phys. Lett. **113**, 093505 (2018)
61. Caponetto, R.; Graziani, S.; Pappalardo, F.L.; Sapuppo, F.: Experimental characterization of ionic polymermetal composite as a novel fractional order element. Adv. Math. Phys. 1–10, Article ID: 953695 (2013)
62. Tu, S.; Jiang, Q.; Zhang, X.; Alshareef, H.N.: Solid state mxene based electrostatic fractional capacitors. Appl. Phys. Lett. **114**(23), 232903 (2019)
63. Vinagre, B.M.; Podlubny, I.; Hernandez, A.; Feliu, V.: Some approximations of fractional order operators used in control theory and applications. Fract. Calc. Appl. Anal. **3**(3), 945–950 (2000)
64. El-Khazali, R.: On the biquadratic approximation of fractional order Laplacian operators. Analog Integr. Circuits Signal Process. **82**(3), 503–517 (2015)
65. Xue, D.Y.; Zhao, C.; Chen, Y.Q.: A modified approximation method of fractional order system. In: IEEE Conference on Mechatronics and Automation, pp. 1043–1048, (2006)
66. Muñiz-Montero, C.; García-Jiménez, L.V.; Sánchez-Gaspariano, L.A.; Sánchez-López, C.; Gonzalez-Diaz, V.R.; Tlelo-Cuautle, E.: New alternatives for analog implementation of fractional-order integrators, differentiators and PID controllers based on integer-order integrators. Nonlinear Dyn. **90**(1), 241–256 (2017)
67. Radwan, A.G.; Soliman, A.M.; Elwakil, A.S.; Sedeek, A.: On the stability of linear systems with fractional-order elements. Chaos Solitons Fractals **40**(5), 2317–2328 (2009)
68. Tlelo-Cuautle, E.; Sanabria-Borbón, A.C.: Optimizing operational amplifiers by evolutionary algorithms and gm/Id method. Int. J. Electron. **103**(10), 1665–1684 (2016)
69. Yan, Z.; Mak, P.; Martins, R.P.: Double recycling technique for folded-cascode OTA. Analog Integr. Circuits Signal Processing. **71**, 137–141 (2012)
70. Joshi, A.; Shrimali, H.; Sharma, S.K.: Systematic design approach for a gain boosted telescopic OTA with cross-coupled capacitor. IET Circuits Devices Syst. **11**(3), 225–231 (2017)
71. Dammak, H.D.; Bensalem, S.; Zouari, S.; Loulou, M.: Design of folded cascode OTA in different regions of operation through gm/ID methodology. Int. J. Electr. Comput. Eng. **2**(9), 1741–1746 (2008)
72. Tsukutani, T.; Sumi, Y.; Fukui, Y.: Electronically controlled current-mode oscillators using MO-OTAs and grounded capacitors. Frequenz **60**(11–12), 220–223 (2006)
73. Adhikary, A.; Sen, P.; Sen, S.; Biswas, K.: Design and performance study of dynamic fractors in any of the four quadrants. Circuits Syst. Signal Process. **35**(6), 1909–1932 (2016)

Journal of Polymer Research (2020) 27:296
https://doi.org/10.1007/s10965-020-02236-9

ORIGINAL PAPER

Electrospun pH responsive poly (acrylic acid-co- acrylamide) hydrogel nanofibrous mats for drug delivery

Preeti Gupta[1] · Roli Purwar[1]

Received: 26 March 2020 / Accepted: 26 May 2020
© The Polymer Society, Taipei 2020

Abstract

In the present study hydrogel nanofibrous mats are prepared through electrospinning of poly (acrylamide-co-acrylic acid) dope solution having polyethylene glycol as cross-linker. The electrospun mats are heat treated at 150 °C to induce esterification crosslinking reaction. The mats are characterized for their structural, morphological and thermal behaviour. The average diameter of hydrogel nanofibers is found to be 275 ± 94.5 nm. The swelling behaviour of nanofibrous mats is evaluated in the phosphate buffer solutions of different pHs in the range of 2 to 8.5 and is found to be highest at neutral pH. The nanofibrous mats are loaded with amoxicillin drug for antimicrobial activity and drug release behaviour. The drug loaded mats form a clear inhibition zone with both Gram-positive bacteria *S. aureus* and Gram-negative bacteria *E. coli* in disc diffusion test. Drug is released at all the pHs (2 to 8.5) and maximum release is observed at neutral pH. The kinetics of drug release follows Fickian diffusion ($n < 0.5$) at tested pH.

Keywords Hydrogel nanofibers · Drug release · Swelling behaviour · Poly (acrylic acid-co-acrylamide) · Fickian diffusion

Introduction

Medicated hydrogels containing drugs/antibiotic, reduce the bacterial infection at wound site have built their interest to ameliorate the currently available wound healing processes [1–5]. Main features of hydrogels like porosity and swelling in aqueous media [6] can easily be modulated by regulating the closeness of cross-links in the hydrogel mats [7]. Porous framework also allows drug loading into the hydrogel mats and its subsequent exoneration at a rate relying on the diffusivity of the micro or macro molecules through the hydrogel network [8].

Electrospun nanofibrous mats have high surface to volume ratio, porous structure, good mechanical strength, more flexibility as compared to their film counterparts [9]. Inverse relation of the swelling kinetics of hydrogels films with the square of their dimensions suggests that the rate of volume change can be raised by altering the morphology of the hydrogel material from films to nanofibrous form. Large surface area of hydrogel nanofibrous mats permits rapid discharge of antibiotics or growth factors at wound site [10, 11]. The highly porous framework of nanofibrous mats able do rapid body fluids absorption and waste diffusion. These alluring properties of hydrogel nanofibrous mats make them a potent candidate for medicated wound dressing material [12].

Recently, efforts have been made on designing the framework and composition for generating advance nanofibers-rooted hydrogel having enhanced mechanical strength and multifunctional operation behaviour. Electrospun hydrogel nanofibers from poly (styrene-co-maleic anhydride) [13], poly (styrene-co-maleic sodium anhydride)/cellulose [14], multiblock poly ester urethane [8], PVA/Hyaluronic acid [15], polyvinyl alcohol/polyacrylic acid/MWCNT [16], Chitosan-g-PNIPAAM/Polyethylene oxide [17], poly aspartic acid [11], poly(vinyl caprolactam-co-hydroxymethyl acrylamide) [18] and Polyvinylpyrrolidone [19] have been explored for applications like tissue engineering, drug delivery etc. An important shortcoming associated with the several hydrogel nanofibers mats was their easy deformation in aqueous media i.e. an unstable fibre morphology in aqueous media makes them non suitable for load bearing applications.

Stimuli responsive hydrogel especially pH, can be formed using polyacrylic acid having activated on/off switch in acidic

✉ Roli Purwar
roli.purwar@dtu.ac.in

[1] Discipline of Polymer Science and Chemical Technology, Department of Applied Chemistry, Delhi Technological University, Shahabad Daulatpur, Bawana Road, Delhi 110042, India

Published online: 05 September 2020

media. The major shortcoming associated with polyacrylic acid based hydrogel nanofibers is the immense deformation of individual fibers when they are dipped in water because of the presence of non-rigid back bone chain structure [13]. The morphology of the hydrogel nanofibers during water immersion/drying cycles can be maintained as far as possible by employing the concept of copolymerization. Hydrogel made up of polyacrylamide can have the ability of imbibition of water up to several times of its own weight. Hydrogel films made up of copolymers of polyacrylic acid (AAc) with polyacrylamide (AAm) derivatives have been reported for several different drug release [5, 16–18, 20]. Therefore, if we can make nanofibers of poly (acrylic acid-co-acrylamide) hydrogel, it may have high absorption capacity for extrudate from wound site as compared to films. Large surface area of hydrogel nanofibrous mats enables pH responsive release of drug at wound site. It will also provide moist environment for faster healing.

As per the information we have, we are first one for reporting the preparation of hydrogel nanofibers of copolymer of acrylic acid and acrylamide using PEG (polyethylene glycol), as a cross linker in two step process. In step one the dope solution of poly (acrylic acid-co-acrylamide) and PEG is electrospun to form nanofibrous mats. These mats are subjected to heat treatment to induced crosslinking. The obtained hydrogel nanofibrous mats (NFHGs) are characterised for their structural, morphological, and thermal properties. The porosity, swelling kinetics and drug release behaviour are evaluated.

Experimental

Materials and methods

Materials

Ammonium per sulphate (APS) and Acrylic acid (AAc) from Central drug house (Delhi, India).Acrylamide (AAm) of Reagent grade from Sisco Research Laboratories (Mumbai, India), Polyethylene glycol (PEG 6000) from LOBA CHEMI (Mumbai, India), Amoxicillin from HIMEDIA Laboratories Pvt. (India) Ltd., Cotton fabric (139 g/m^2) were employed as received. Distilled water was used for carrying out all the experiments.

Methods

Development of copolymer of poly (acrylic acid-co-acrylamide) Copolymer of poly (acrylic acid-co-acrylamide) was developed using solution polymerisation method. In a 1 M aqueous solution of acrylamide a 10 ml mixture of 0.1 mol acrylic acid and 0.076 g ammonium per sulphate was added drop wise and with constant shaking. The reaction was conducted for about 3 h at a temperature of 70 °C. The copolymer precipitated out and washed with a methanol water mixture of ratio 1:1. The copolymer was filtered and kept in a vacuum oven at 60 °C for about 10 h to make it moisture free.

Development of hydrogel nanofibrous mat The dope solution for electro spinning was prepared by dissolving poly (acrylic acid-co-acrylamide) copolymer 8% (*w*/*v*) in distilled water with continuous shaking at normal temperature. Polyethylene glycol (2% *w*/w) was added in the dope solution. The copolymer dope solution was electrospun using a voltage of 23 kV, tip to collector distance of 15 cm and feed rate of 3.5–4 μLmin^{-1} on a drum collector. Cotton fabric is used for collecting the nanofibrous mats. The nanofibrous mats were converted into hydrogel form by placing it in an oven at 150 °C for 10 min and the obtained hydrogel nanofibrous mats were dried in a vacuum oven at 80 °C for 10–12 h. Schematic diagram for development of hydrogel nanofibrous mats is shown in Fig. 1.

Characterization

Morphology of the hydrogel nanofibrous mats was studied using Scanning Electron Microscope, SEM (model no. S-3700 N, Hitachi, Germany). For this the samples were first made electrically conductive by gold-sputter coating and then fixed using adhesive tape in the sample stage. Diameter distribution of the nanofibers was obtained by analysing the SEM images for 50 fibres using J-image software. The crosslinking of nanofibrous hydrogel was analysed using FTIR (Thermoscientific Nicolet 380 spectrophotometer, USA) in ATIR mode. Scanning was done in the frequency range of 500 - 4000 cm^{-1} and resolution of 4 cm^{-1}.

The thermal properties were tested using Perkin Elmer TGA 4000 instrument under nitrogen atmosphere with scanning rate 10 °C/min and temperature range 30 °C to 700 °C. The differential scanning calorimetry of the nanofibrous mats was carried out using Perkin Elmer DSC8000 instrument in the temperature range of 30 °C to 160 °C and scanning rate 10 °C /min under nitrogen atmosphere.

Porosity of the nanofibrous mat was evaluated using liquid displacement method. Here solvent used was n-hexane as it shows unhindered motion through interlinked pores and have no effect on the morphology of nanofibrous mat. For measurement known amount of solvent, here 7 ml, was taken in a measuring cylinder and then the sample of specific size was kept inside it for 10 min. The volume of the solvent was measured after immersion of sample and after removal of hexane-soaked sample. Following formula is used to calculate the porosity of nanofibrous mats from the volumes obtained

$$\%\text{Porosity} = (V0-V2)*100/(V1-V2) \quad (1)$$

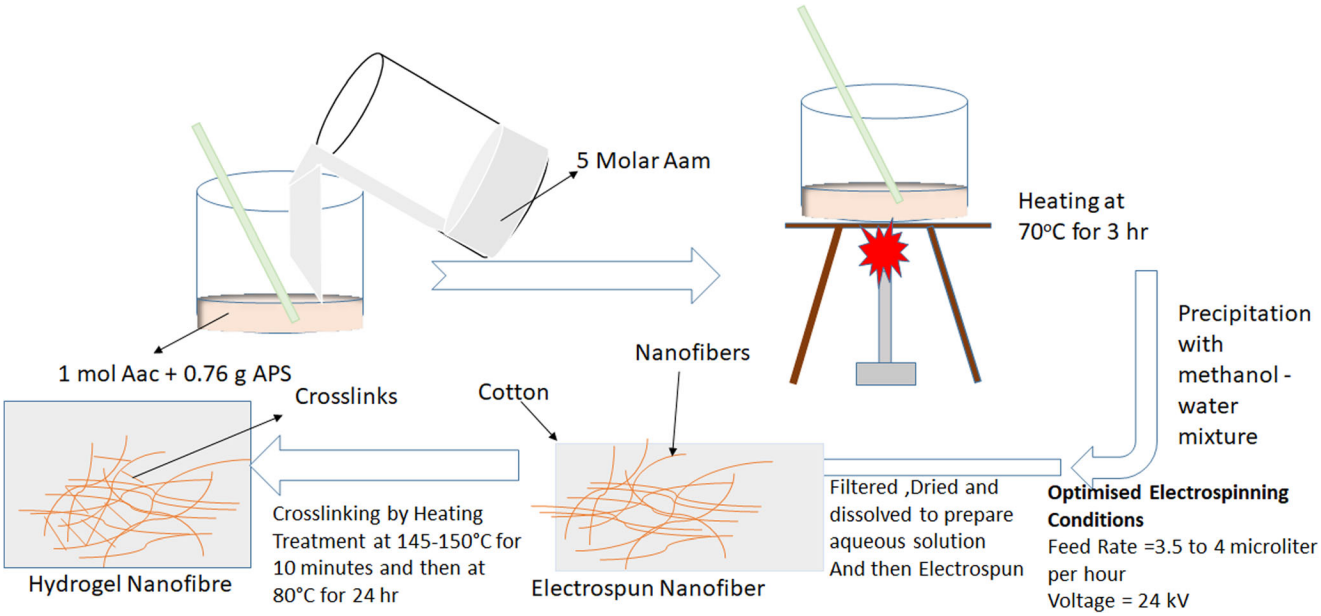


Fig. 1 Schematic diagram for preparation of hydrogel nanofibrous mat

Where V0 = Volume of n-hexane taken, V1 = Volume of solvent after dipping the mat, V2 = Volume of solvent after removing hexane soaked mat.

The porosity of hydrogel nanofibrous mat was found to be 75% while porosity in case of hydrogel nanofibers of poly(-vinyl alcohol)/poly (acrylic acid) was reported to be 40% [21].

Swelling analysis

Swelling behaviour of hydrogel nanofibrous mats was evaluated as described by Anuja and Purwar [22]. In brief, the 3X4 mm^2 samples of hydrogel nanofibers were plunged in the PBS buffer solutions of pH 2, 4.5, 5.5, 7, 8.5 at 25 °C and their weights were taken after regular time interval till equilibrium is reached by wiping out the excess of surface water with the help of tissue paper. The experiment was performed in triplicate and their mean was taken for analysis. The water present in hydrogel nanofibrous mat may be expressed as follows

Percent isothermal swelling degree (SDt)

$$= \frac{wt - w0}{wt} \times 100 \quad (2)$$

here w0 = weight of hydrogel nanofibrous mat at t=0
wt= weight of hydrogel nanofibrous mat at time t

$$\text{Normalize swelling ratio}(\alpha) = \frac{SDt}{SDeq} \quad (3)$$

where SD_{eq} =equilibrium swelling degree at a certain temperature and pH.

Swelling kinetics for hydrogel nanofibrous samples were analysed using different kinetics models like Peppas, first order and second order. The model for which value of correlation coefficient (R^2) is highest amongst all, explains the kinetics swelling of hydrogel or its nanofibrous form [19, 20, 23, 25]. The mathematical formulation of Peppas model is interpreted as

$$\alpha \text{ Or } M_t/M_\infty = kt^n \quad (4)$$

where, α is degree of normalized swelling, n is swelling exponent, showing mode of transport mechanism of the penetrate, k is constant of hydrogel, t is swelling time.

First order kinetic model is interpreted as

$$\alpha = \left(1 - Ae^{-kt}\right) \quad (5)$$

where, A is the pre exponential factor.

Second order kinetic model is interpreted as

$$t/M_t = \left(1/kM_\infty^2\right) + (1/M_\infty)t \quad (6)$$

where, M_∞ is equilibrium weight of hydrogel nanofibrous mat and M_t is weight of nanofibrous mat at time t.

Antibacterial test

AATCC -30 method was used to test the antibacterial activity of amoxicillin drug loaded hydrogel nanofibrous mats. For this study, the model bacteria used were *E. coli.* (Gram negative) and *S. aureus*. (Gram positive). The drug loaded hydrogel nanofibrous mat was cut into discs and sterilized by placing in a laminar air flow under UV light

for about 24 h. The model bacteria were cultured in the agar plates and the discs of drug loaded mat were kept at the centre of these cultured agar plates. Close contact of the disc with the surface should be ensured by pressing it lightly. These plates were kept in an incubator at 37 °C for about 24 h. After 24 h, the size of the inhibition zone is measured in order to predict the efficacy of drug loaded mats. In all cases the experiment were carried out in triplicate.

Drug release kinetics study

The kinetics of drug release from hydrogel nanofibrous mats at different pHs was studied by using Carry 300 UV-Visible spectrophotometer (Agilent Technologies). Amoxicillin solutions with concentrations ranging from 0.0 to 0.1 mg per ml were used to prepare a calibration curve at 273 nm [26]. For studying release kinetics, the dried nanofibrous mats of 50 mg were kept in the test tubes having 10 mL aqueous solutions of amoxicillin (1 mg per ml) for 48 h at normal temperature. After completion of 48 h, the samples were taken out from the solutions and dried. Difference in the concentrations of amoxicillin solution before and after drug loading give the amount of drug loaded in the nanofibrous mat. Here, this amount is found to be 0.589 mg. This value (0.589 mg) is taken as the maximum amount of the drug that can be come out from the dressings (M_∞) for the calculations. The dried drug loaded samples were then placed separately in the 10 mL solutions of phosphate buffer of pH 2(highly acidic), pH 5.5 (acidic), pH 7(neutral) and pH 8.5(basic).To find out the amount of drug released at different time intervals, 1 mL aliquot was taken out from each buffer solution and the same amount of fresh dissolution medium (buffer) was added simultaneously, after each withdrawal. All the drug release experiments were performed for 48 h. From the standard curve, cumulative drug release was determined. All the experiments were repeated thrice, and their mean values were taken for analysis. The obtained data were fitted in Peppas and Higuchi kinetic models to determine the value of release constant [24]. The mathematical formulation of Peppas model is given as

$$M_t/M_\infty = kt^n \quad (7)$$

where M_t is cumulative amount of the drug released at time t, M_∞ is cumulative amount of the drug released at infinite time; k is the constant related with the structural and geometric characteristics of the nanofibrous mats; and n is the release exponent, an indicator of drug release mechanism.

Higuchi model can be interpreted as

$$M_t = M_\infty + k_H t^{1/2} \quad (8)$$

where k_H is Higuchi constant, an indicator of characteristics of a hydrogel system.

The model for which value of correlation coefficient (R^2) is highest amongst all, was considered best for the drug release mechanism.

Statistical analysis

All the results are taken as mean ± SD. One way ANOVA was done to assess the experimental data at the statistical significance level of $p < 0.05$.

Results and discussion

Morphology of hydrogel nanofibrous mats

SEM image of the hydrogel nanofibrous mat is shown in Fig. 2. Bead less nanofibrous morphology was obtained at 23 kV voltage however, there appeared a large variation in the fibre diameter. The average diameter of the nanofibers mats was found to be 275 nm ± 94.5. This is attributed due to change in viscosity of the solution with passage of time. The average diameter and bead less structure of nanofibers are dependent on concentration of polymer in dope, molecular weight of the polymer, process parameters of electrospun

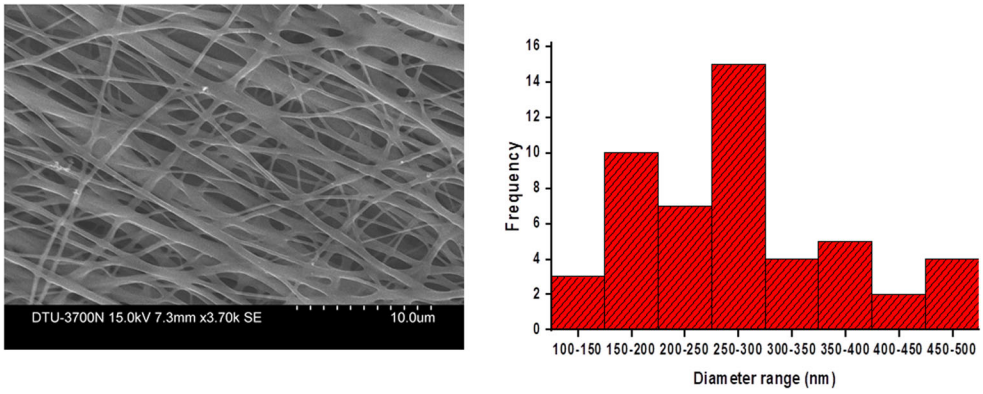


Fig. 2 **a** SEM Image of nanofibrous mat (**b**) J Image analysis for fiber diameter

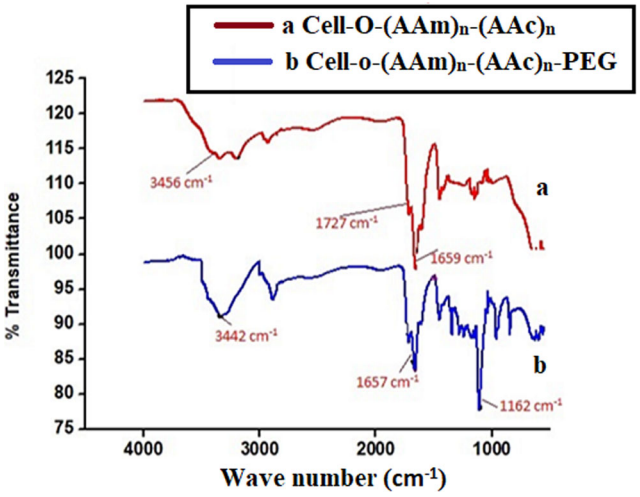


Fig. 3 FTIR curves for uncross-linked and cross-linked nanofibrous mats

machines. Similar diameter range was reported for hydrogel nanofibers made from CTS-g PNIPAAm/poly(ethylene oxide) [17] and PVA [27] . The diameter of poly (acrylamide-co-maleic acid) hydrogel nanofibers is lower i.e. 120 nm [28]. The average diameter of Chitosan/PLA/PEG and chitosan/PVA hydrogel nanofibers was reported to be 341 and 430 nm respectively.

Structural properties

In Fig. 3 FTIR spectra of non-cross linked and cross-linked nanofibrous mats are shown. Figure 3a clearly reveals the characteristic peaks of copolymer of poly (acrylic acid-co-polyacrylamide).In this spectra a peak at 3456 cm^{-1} is due to N-H stretching, at 1727 cm^{-1} it is for C=O (C=O of acidic group) stretching and at 1659 cm^{-1} it is for C=O (C=O of acrylamide group) stretching. In Fig. 3b, an extra peak at 1167 cm^{-1} is for C-O-C stretching showing the presence of PEG crosslinking in hydrogel. Other characteristic peaks of poly (acrylic acid-co-acrylamide) cross-linked with PEG like N-H stretching at 3442 cm^{-1}, C=O (C=O of acrylamide group) stretching at 1658 cm^{-1} and C=O (C=O of acidic group) stretching at 1727 cm^{-1}also appear in it. The reaction mechanism for the formulation of hydrogel nanofibrous mat is indicated in Fig. 4.These results were in close agreement with the FTIR results reported by Purwar et al where they examined the cotton grafted poly (acrylamide-co-acrylic acid) hydrogel [29].

Thermal properties

The thermal properties of cross-linked and non-cross-linked nanofibrous mats were analysed through DSC and TGA. Figure 5 shows the DSC thermogram of cross-linked and non-cross-linked nanofibrous mats. The non crosslinked sample and cross-linked nanofibrous mats show endothermic peaks at 73.22 °C and 64.75 °C due to loss of water [30]. The values of enthalpy are 0.694 Jg^{-1} and 38.97 Jg^{-1} at endothermic peak of 73.22 °C and 64.75 °C respectively. The higher enthalpy value for cross-linked nanofibrous mats as compared to that of uncross-linked sample suggests that the nanofibrous hydrogel has more capacity to hold water as compared to its non-cross-linked counterpart.

The thermal degradation behaviour of non-cross-linked and cross-linked nanofibrous mats is shown in Fig. 6. Both the samples show three step degradation behaviour. Initially up to 100 °C the mass loss is due to water content. The cross-linked sample shows higher thermal stability as compared to uncross-linked sample above 400 °C.

Swelling behaviour

The swelling behaviour of hydrogel nanofibrous mat with time was studied at different pH in the range of 2 to 8.5 and the results are shown in Fig. 7. It was observed that nanofibers attained its maximum swelling percentage within 10 h at tested pH. However, continuous swelling was observed under pH 7 up to 50 h. The percent swelling increases with increase in pH up to 7, further there is no increase in swelling percentage. The maximum swelling percent of 430% was achieved at

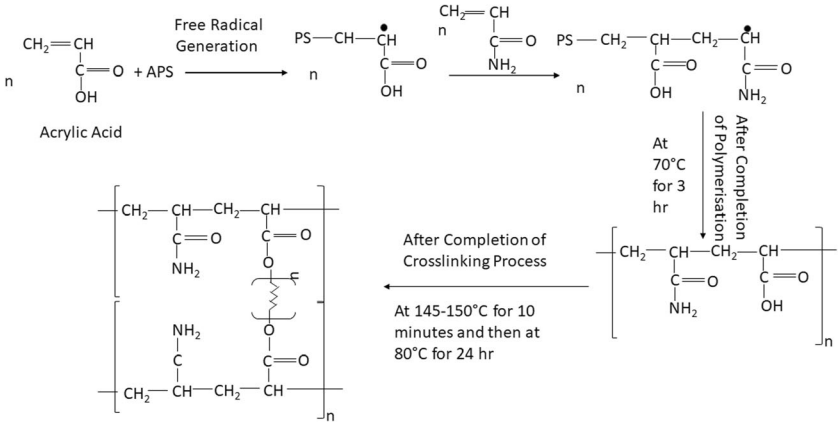


Fig. 4 Reaction mechanism for formation of hydrogel

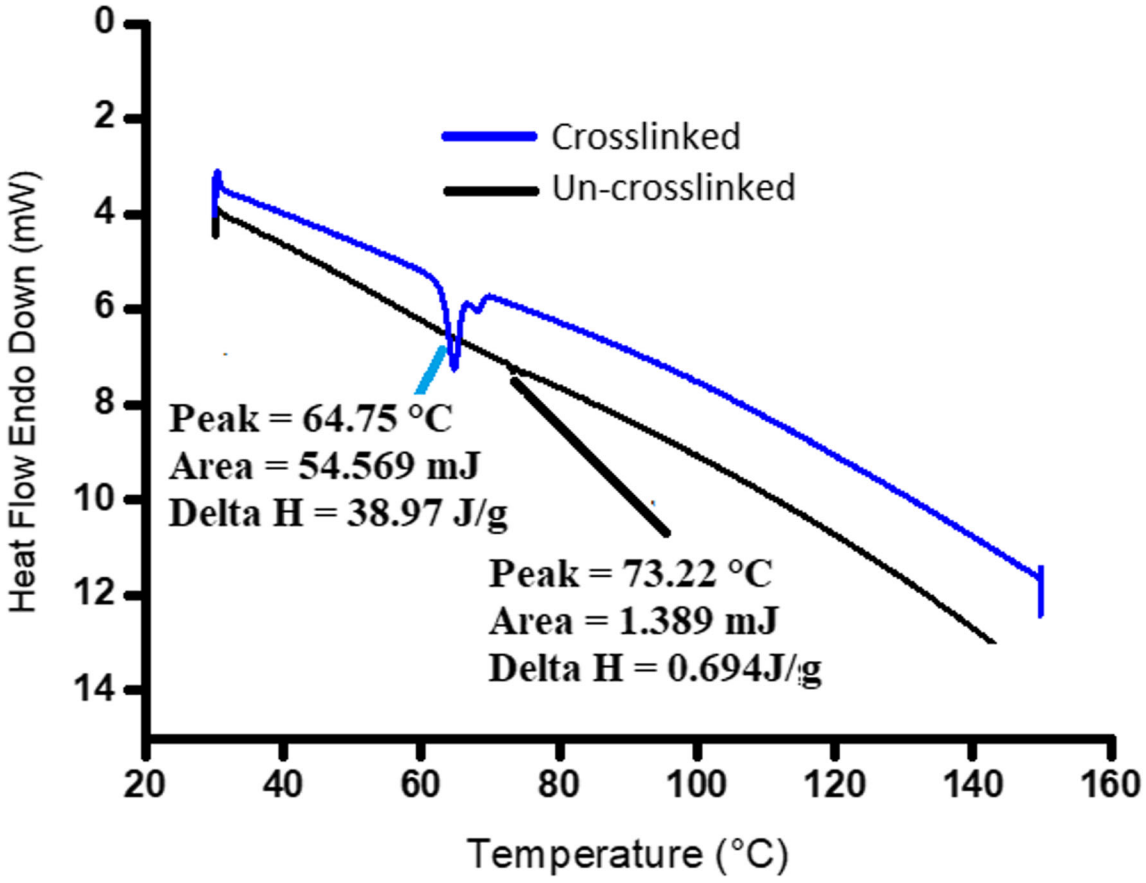


Fig. 5 DSC thermogram of uncross-linked and cross-linked nanofibrous mats

pH 7.This swelling behaviour Poly (acrylic acid-co-acrylamide) hydrogel nanofibrous mat can be attributed by the fact that at lower pH i.e. below 4.5 all the pendant groups exist in unionized form and some hydrogen bonds may exist between carboxylic acid and amino groups. Formation of hydrogen bonds reduces the available space for holding water molecules resulting in low percentage swelling. At high pH (i.e. above than the pKa of acrylic acid) the carboxylic groups ionise to generate carboxylate ions, which repel each other due to the presence of negatively charged centres and solvation of charged group, resulting in more pore size and higher percent swelling. In alkaline pH (8.5) a decline in percent swelling is because of the dissociation of carboxylic group in alkaline medium. Due to which the physical forces existing between poly acrylic acid and poly acrylamide became weak, hence pores got diminished. Similar results were observed in our previous study where poly (acrylamide-co-acrylic acid) hydrogel cross-linked with polyethylene glycol grafted on cotton fabric. A major observation was found that the nanofibrous mat achieved equilibrium swelling in 10 h. In case of grafted hydrogel films the equilibrium swelling was achieved after 20 h, however the amount of maximum swelling percentage was achieved 800% [22]. The swelling percentage of Polyacrylic acid nanofibers cross-linked with β-cyclodextrin was reported 1800% at pH 7. Jin and Hsieh [21] have compared swelling behaviour of hydrogel films and nanofibrous mat prepared with Poly vinyl alcohol/Poly acrylic acid blend and reported that percent swelling increases with rise in pH. The mass swelling of films were found higher than that of nanofibrous mat. The nanofibrous hydrogel attained equilibrium within 10 min as compared to 30 min in the film form.

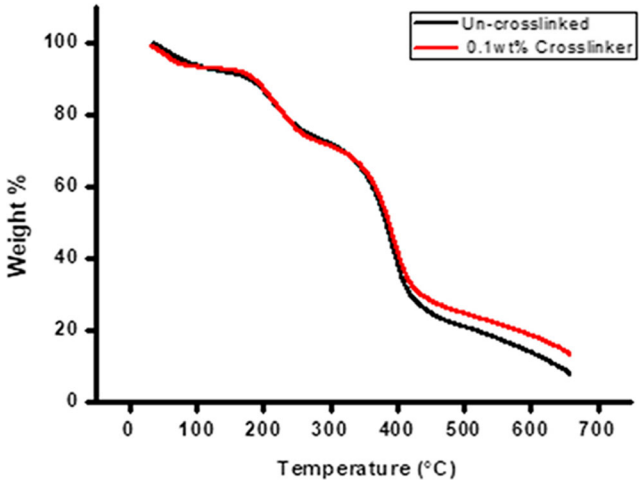


Fig. 6 TGA thermogram of un-cross-linked and cross-linked nanofibrous mats

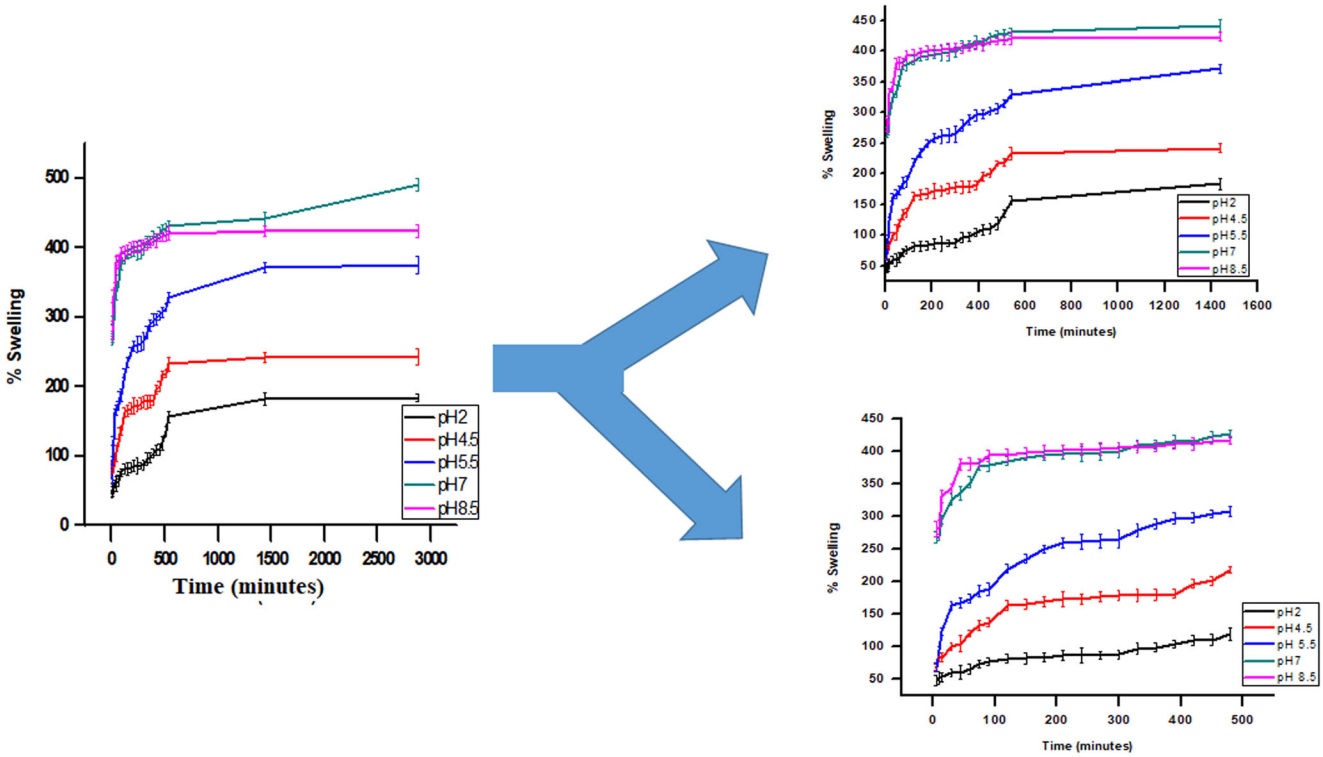


Fig. 7 Swelling behavior of nanofibrous mat at different pH

Swelling kinetics

Swelling kinetics of the hydrogel nanofibrous mats was studied by fitting the swelling data in different kinetic models namely Peppas model, first order and second order kinetics equations. The plots obtained by these model/equations are shown in Figs. 8 (a) to (c). Out of these the plot having maximum linearity was considered as best fit kinetic model. Kinetic constants for swelling at different pHs using various kinetic models have been summed up in Table 1.Value of correlation coefficient (R^2) clearly shows that swelling follows second order kinetics at all the pHs taken for study (2, 5.5, 7, 8.5).

Literature review shows that swelling kinetics of polymer hydrogel may be either diffusion controlled (Fickian) or relaxation controlled (non- Fickian).The mechanism of water insertion into the hydrogel nanofibrous mats is governed by diffusion of water molecules into the gel matrix and subsequent relaxation of polymeric chain of hydrogel nanofibers. Literature also revealed first order swelling kinetics for poly (acrylic acid) hydrogel at all the pHs studied while second order swelling kinetics for poly (acrylamide-co-itaconic acid) /chitosan and chitosan-g-PAAm hydrogel [31].

In this study swelling kinetics is of second order at all the pHs studied i.e. follows relaxation controlled mechanism. It suggests swelling rate varies directly with the square of water

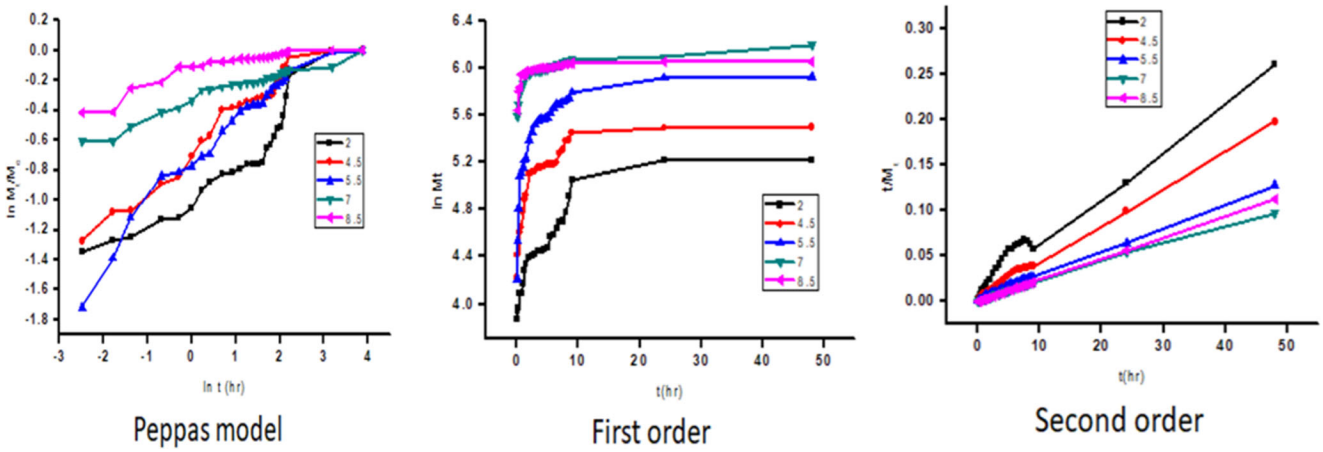


Fig. 8 Swelling kinetics of nanofibrous mat

Table 1 Kinetic constants for swelling at different pH

	Peppas model			First order model		Second order model	
pH	$(R^2)^a$	n^b	k^c	$(R^2)^a$	k^c	$(R^2)^a$	k^c
2	0.89	0.23	0.11	0.57	0.027	0.96	0.0017
4.5	0.955	0.23	0.21	0.32	0.02	0.99	0.0032
5.5	0.95	0.27	0.164	0.30	0.029	0.99	0.0018
7	0.98	0.094	0.451	0.399	0.009	0.99	0.0027
8.5	0.86	0.066	0.7	0.19	0.0046	1	0.013

[a] Correlation coefficient [b] Release exponent [c] Constant of hydrogels

content retained in the nanofibrous mat before attaining equilibrium. With passage of time, swelling rate decreases rapidly because of its dependence upon the osmotic pressure difference. Swelling kinetic of poly (acrylic acid -co- acrylamide) hydrogel films emulate first order kinetics at pH 5.5 and 7.0 while at pH 8.5 it emulates second order kinetics [19].

Antibacterial activity and drug release behaviour

A clear inhibition zone was generated within and around the drug loaded samples in the 24 h incubated agar plates. This shows the bactericidal behaviour of the hydrogel nanofibrous mats. The growth of inhibition zone for both Gram-positive and Gram-negative bacteria were measured. From Fig. 9, the mean diameter of inhibition zone for Gram-positive bacteria *S. aureus* was found around 13.5 mm and for Gram-negative bacteria *E. coli* it was found 12 mm.

The results showed that amoxicillin drug readily comes out from nanofibrous mats and show very good zone of inhibition around the sample. The amoxicillin released from the nanofibrous mats ruined the cell walls of both the bacterial strains [34].

Figure 10 shows the amount of drug discharged as a function of time at different pHs. From the Figure it is evident that a considerable amount of drug is discharged at all the pHs and maximum liberation is observed at pH 7. Up to 68% of drug comes out in first 9 h.

Drug release kinetics

To study drug release kinetics, the drug release data was fitted into various kinetic models like Peppas model, Higuchi model etc. The graphs of Peppas model and Higuchi model are shown in Fig. 10 (a) and (b) respectively. The graph showing highest linearity was considered as best fit model. The kinetic constants for drug release at different pHs for different kinetic models were summarized in Table 2. From this table it is clear that drug release kinetics is best described by Peppas model as the value of correlation coefficient is found to be more than 0.9.

Studies suggest that a drug delivery system with slab geometry and values of release exponent (n) < 0.5, $0.5 < n < 1.0$ or 1.0 corresponds to Fickian diffusion, anomalous transport and case (II) transport (zero order release) respectively. Theophylline release from poly (acrylic acid co acrylamide) hydrogels follows an anomalous kinetics [32] while gentamicin sulphate liberation from poly (acrylamide co acrylic acid) shows anomalous diffusion at low acrylic acid concentration and Fickian diffusion at high acrylic acid concentration [33].

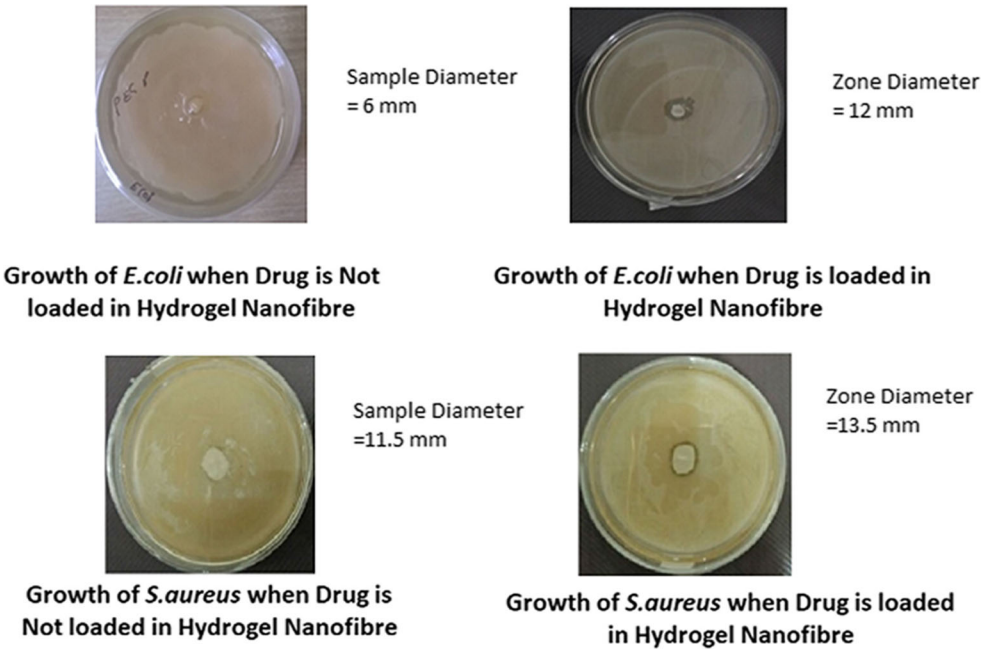


Fig. 9 Antibacterial activity of amoxicillin loaded hydrogel nanofibrous mats

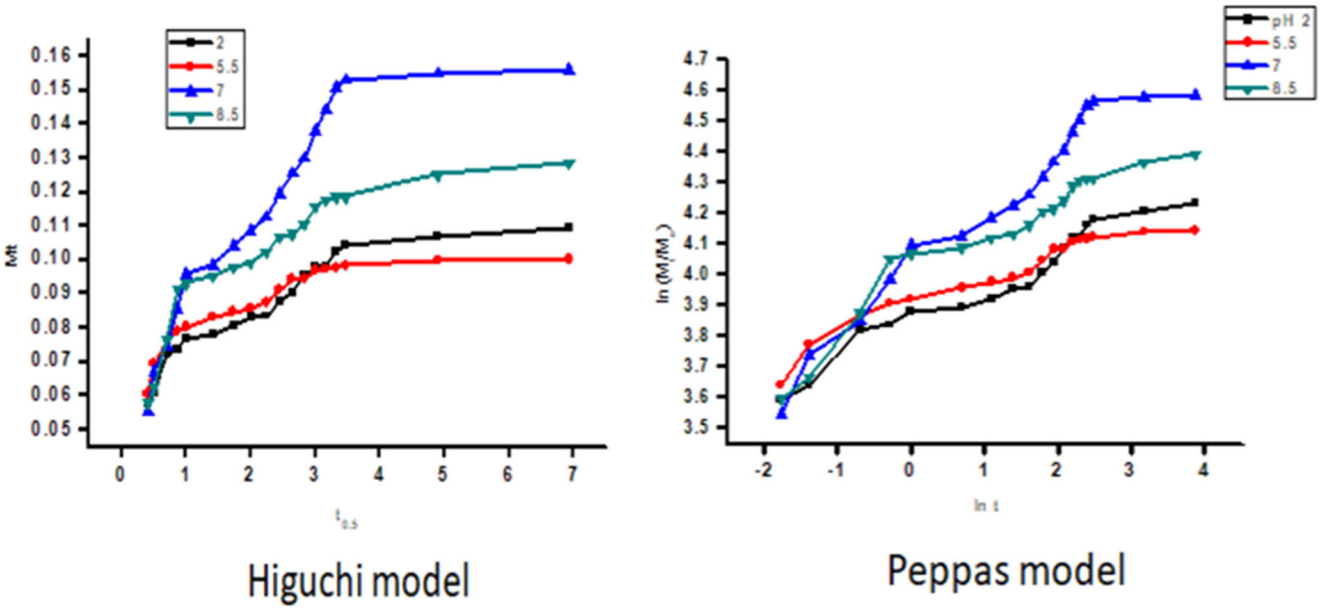


Fig. 10 Drug release kinetics of amoxicillin loaded hydrogel nanofibrous mats

For hydrogel nanofibrous mat release exponent value is below 0.5 at all the pHs which shows Fickian diffusion at all the pHs. It means that the drug diffusion rate is less as compared to the relaxation rate of the hydrogel matrix. The table also indicates that value of k raises with increase in pH up to 7 and then decreases. Since k is related with the geometry of the hydrogel, so it can be assumed that the pore size in the nanofibers hydrogel is maximum at pH 7. This is because of the repulsion between the negatively charged carboxylate groups formed due to the ionization of carboxylic acid groups at this pH. However, above this pH, the physical forces working between poly acrylic acid and poly acrylamide diminished and hence pore size reduces resulting in less drug release.

A lack towards the study of drug release kinetics from hydrogel nanofibers was appeared in previous papers. Yuan et al studied the release of BSA as a model protein from hydrogel nanofibers of CTS-g-PNIPAAm /PEO and reported that BSA release is a diffusion controlled process [17]. Agarwal and Purwar studied the drug release kinetics of hydrogel films of poly(acrylamide- co- acrylic acid) and reported that it fits best in Peppas model i.e. drug release occur by Fickian diffusion mechanism [22].This is also in accordance with the results obtained for hydrogel nanofibers.

Table 2 Kinetic constants for Drug release at different pH

	Peppas model			Higuchi model	
pH	$(R^2)^a$	n^b	k^c	$(R^2)^a$	k^c
2	0.95	0.12	46.26	0.82	0.008
5.5	0.93	0.21	49.74	0.7	0.006
7	0.95	0.28	54.47	0.78	0.016
8.5	0.91	0.15	52.17	0.74	0.010

[a] Correlation coefficient [b] Release exponent [c] Constant of hydrogels

Conclusions

Copolymer of poly(acrylic acid-co-acrylamide) was converted into nanofibrous form using an electrospinning machine with a voltage of 23 kV, tip to collector distance 15 cm and feed rate of 3.5–4 $\mu L min^{-1}$ and nanofibers were cross-linked by heat induced esterification at 150 °C. The obtained fibres have average fibre diameter 275 nm ± 94.5 nm. These mats are highly porous structure with porosity 74%. Swelling of these mats was highest at neutral pH and it follows second order kinetics at all pHs. Half of the swelling is achieved within a period of two hours. The nanofibrous mats were loaded with amoxicillin drug for antimicrobial activity. Formation a clear inhibition zone with both gram positive and gram negative bacteria in disc diffusion test confirmed the loading and releasing of the drug from the hydrogel nanofibers mat. The drug release kinetics follow Fickian diffusion ($n < 0.5$). Initial excessive swelling and a higher initial release rate make the hydrogel nanofibrous mat a potent candidate for making active wound dressings.

Acknowledgements This study was financially supported by UGC (University Grants Commission), under Joint CSIR- UGC (NET)-JRF scheme {F.11-3/2003(SA-I)}.

References

1. Li S, Dong S, Xu W, Tu S, Yan L, Zhao C, Ding J, Chen X (2018) Antibacterial hydrogels. Adv Sci 5:1–17. https://doi.org/10.1002/advs.201700527
2. Sabitha M, Rajiv S (2014) Preparation and characterization of ampicillin-incorporated electrospun polyurethane scaffolds for wound healing and infection control. Polym Eng Sci 55:541–548. https://doi.org/10.1002/pen
3. Jalababu R, Rao KSVK, Rao BS, Reddy KVNS (2020) Dual responsive GG-g-PNPA/PIPAM based novel hydrogels for the controlled release of anti- cancer agent and their swelling and release kinetics. J Polym Res 27:83. https://doi.org/10.1007/s10965-020-02061-0
4. Olad A, Eslamzadeh M, Katiraee F, Mirmohseni A (2020) Evaluation of in vitro anti-fungal properties of allicin loaded ion cross-linked poly (AA-co-AAm)/PVA/Cloisite 15A Nanocomposite hydrogel films as wound dressing materials. J Polym Res 27:100. https://doi.org/10.1007/s10965-020-02072-x
5. Palem RR, Shimoga G, Rao KSVK, Lee SH, Kang TJ (2020) Guar gum graft polymer-based silver nanocomposite hydrogels: synthesis, characterization and its biomedical applications. J Polym Res 27:68. https://doi.org/10.1007/s10965-020-2026-8
6. Uddin MS, Ju J (2020) Effect of crosslinking agents on drug distribution in chitosan hydrogel for targeted drug delivery to treat cancer. J Polym Res 27:81. https://doi.org/10.1007/s10965-020-02059-8
7. Hajikarimi A, Sadeghi M (2020) Free radical synthesis of crosslinking gelatin base poly NVP/acrylic acid hydrogel and nanoclay hydrogel as cephalexin drug deliver. J Polym Res 27:1–20. https://doi.org/10.1007/s10965-020-2020-1
8. Loh XJ, Peh P, Liao S, Sng C, Li J (2010) Controlled drug release from biodegradable thermoresponsive physical hydrogel nanofibers. J Control Release 143:175–182. https://doi.org/10.1016/j.jconrel.2009.12.030
9. Nayak R, Padhye R (2017) Nano Fibres by electro spinning: properties and applications. J Text Eng Fash Technol 2:486–497. https://doi.org/10.15406/jteft.2017.02.00074
10. Chen H, Hsieh Y Lo (2004) Ultrafine hydrogel fibers with dual temperature- and pH-responsive swelling behaviors. J Polym Sci Part A Polym Chem 42:6331–6339. https://doi.org/10.1002/pola.20461
11. Zhang C, Wu S, Qin X (2014) Facile fabrication of novel pH-sensitive poly(aspartic acid) hydrogel by crosslinking nanofibers. Mater Lett 132:393–396. https://doi.org/10.1016/j.matlet.2014.06.031
12. Uppal R, Ramaswamy GN, Arnold C, Goodband R, Wang Y (2011) Hyaluronic acid nanofiber wound dressing — production , characterization , and in vivo behavior. J Biomed Mater Biomater 97B:20–29. https://doi.org/10.1002/jbm.b.31776
13. Tang C, Ye S, Liu H (2007) Electrospinning of poly(styrene-co-maleic anhydride) (SMA) and water-swelling behavior of crosslinked/hydrolyzed SMA hydrogel nanofibers. Polymer 48:4482–4491. https://doi.org/10.1016/j.polymer.2007.05.041
14. Cao S, Liu H (2009) Synthesis of pH-responsive crosslinked poly [styrene- co - (maleic sodium anhydride)] and cellulose composite hydrogel nanofibers by electrospinning. Polym Int 58:545–551. https://doi.org/10.1002/pi.2565
15. Kim K, Akada Y, Kai W et al (2011) Cells attachment property of PVA hydrogel Nanofibers incorporating hyaluronic acid for tissue engineering. J Biomater Nanobiotechnol 2:353–360. https://doi.org/10.4236/jbnb.2011.24044
16. Yun J, Im JS, Lee YS, Kim H Il (2011) Electro-responsive transdermal drug delivery behavior of PVA/PAA/MWCNT nanofibers. Eur Polym J 47:1893–1902. https://doi.org/10.1016/j.eurpolymj.2011.07.024
17. Yuan H, Li B, Liang K, Lou X, Zhang Y (2014) Regulating drug release from pH- and temperature-responsive electrospun CTS-g-PNIPAAm/poly(ethylene oxide) hydrogel nanofibers. Biomed Mater 9:055001. https://doi.org/10.1088/1748-6041/9/5/055001
18. González E, Frey MW (2017) Synthesis, characterization and electrospinning of poly(vinyl caprolactam-co-hydroxymethyl acrylamide) to create stimuli-responsive nanofibers. Polymer 108:154–162. https://doi.org/10.1016/j.polymer.2016.11.053
19. Lubasova D, Niu H, Zhao X, Lin T (2015) Hydrogel properties of electrospun polyvinylpyrrolidone and polyvinylpyrrolidone/poly(-acrylic acid) blend nanofibers. RSC Adv 5:54481–54487. https://doi.org/10.1039/c5ra07514a
20. Yang Z, Peng H, Wang W, Liu T (2010) Crystallization behavior of poly(ε-caprolactone)/layered double hydroxide nanocomposites. J Appl Polym Sci 116:2658–2667. https://doi.org/10.1002/app
21. Jin X, Hsieh Y Lo (2005) PH-responsive swelling behavior of poly(vinyl alcohol)/poly(acrylic acid) bi-component fibrous hydrogel membranes. Polymer 46:5149–5160. https://doi.org/10.1016/j.polymer.2005.04.066
22. Agrawal A, Purwar R (2018) Swelling and drug release kinetics of composite wound dressing. Indian J Fibre Text Res 43:104–111
23. Serra L, Doménech J, Peppas NA (2006) Drug transport mechanisms and release kinetics from molecularly designed poly(acrylic acid-g-ethylene glycol) hydrogels. Biomaterials 27:5440–5451. https://doi.org/10.1016/j.biomaterials.2006.06.011
24. Adnadjevic B, Jovanovic J (2009) A comparative kinetics study of isothermal drug release from poly(acrylic acid) and poly(acrylic-co-methacrylic acid) hydrogels. Colloids Surfaces B Biointerfaces 69:31–42. https://doi.org/10.1016/j.colsurfb.2008.10.018
25. Martínez-Ruvalcaba A, Sánchez-Díaz JC, Becerra F et al (2009) Swelling characterization and drug delivery kinetics of polyacrylamide-co-itaconic acid/chitosan hydrogels. Express Polym Lett 3:25–32. https://doi.org/10.3144/expresspolymlett.2009.5
26. Liu Z, Lu W, Qian L, Zhang X, Zeng P, Pan J (2005) In vitro and in vivo studies on mucoadhesive microspheres of amoxicillin. J Control Release 102:135–144. https://doi.org/10.1016/j.jconrel.2004.06.022
27. Çay A, Miraftab M (2013) Properties of electrospun poly(vinyl alcohol) hydrogel nanofibers crosslinked with 1,2,3,4-butanetetracarboxylic acid. J Appl Polym Sci 129:3140–3149. https://doi.org/10.1002/app.39036
28. Liu H, Zhen M, Wu R (2007) Ionic-strength- and pH-responsive poly[acrylamide-co-(maleic acid)] hydrogel nanofibers. Macromol Chem Phys 208:874–880. https://doi.org/10.1002/macp.200600632
29. Purwar R, Rajput P, Srivastava CM (2014) Composite wound dressing for drug release. Fibers Polym 15:1422–1428. https://doi.org/10.1007/s12221-014-1422-2
30. Nesrinne S, Djamel A (2017) Synthesis, characterization and rheological behavior of pH sensitive poly(acrylamide-co-acrylic acid) hydrogels. Arab J Chem 10:539–547. https://doi.org/10.1016/j.arabjc.2013.11.027
31. Pourjavadi A, Mahdavinia GR (2006) Superabsorbency, pH-sensitivity and swelling kinetics of partially hydrolyzed chitosan-g-poly(acrylamide) hydrogels. Turkish J Chem 30:595–608
32. Mendiza E, Puig J, Katime I, Novoa R (1999) Theophylline release from poly (acrylic acid-co-acrylamide). Polym Test 18:559–566.https://doi.org/10.1016/s0142-9418(98)00054-3
33. Thakur A, Wanchoo RK, Singh P (2011) Hydrogels of poly(acrylamide-co-acrylic acid):in-vitro study on release of gentamicin sulfate. Chem Biochem Eng Q 25:471–482 https://hrcak.srce.hr/76440
34. Akhavan BJ, Vijhani P (2019) Amoxicillin. In StatPearls [Internet]. Publishing, StatPearls

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Proceedings of the Second International Conference on Inventive Research in Computing Applications (ICIRCA-2020)
IEEE Xplore Part Number: CFP20N67-ART; ISBN: 978-1-7281-5374-2

Email Spam Detection Using Machine Learning Algorithms

Nikhil Kumar
Computer Science and Engineering Department
Delhi Technological University
New Delhi, India
nikhilkmr445@gmail.com

Sanket Sonowal
Computer Science and Engineering Department
Delhi Technological University
New Delhi, India
sanketsonowal@gmail.com

Nishant
Computer Science and Engineering Department
Delhi Technological University
New Delhi, India
nishantyadav420.ny@gmail.com

***Abstract*—- Email Spam has become a major problem nowadays, with Rapid growth of internet users, Email spams is also increasing. People are using them for illegal and unethical conducts, phishing and fraud. Sending malicious link through spam emails which can harm our system and can also seek in into your system. Creating a fake profile and email account is much easy for the spammers, they pretend like a genuine person in their spam emails, these spammers target those peoples who are not aware about these frauds. So, it is needed to Identify those spam mails which are fraud, this project will identify those spam by using techniques of machine learning, this paper will discuss the machine learning algorithms and apply all these algorithm on our data sets and best algorithm is selected for the email spam detection having best precision and accuracy .**

Keywords: Machine learning, Naïve Bayes, support vector machine-nearest neighbor, random forest, bagging, boosting, neural networks.

I. INTRODUCTION

Email or electronic mail spam refers to the "using of email to send unsolicited emails or advertising emails to a group of recipients. Unsolicited emails mean the recipient has not granted permission for receiving those emails. "The popularity of using spam emails is increasing since last decade. Spam has become a big misfortune on the internet. Spam is a waste of storage, time and message speed. Automatic email filtering may be the most effective method of detecting spam but nowadays spammers can easily bypass all these spam filtering applications easily. Several years ago, most of the spam can be blocked manually coming from certain email addresses. Machine learning approach will be used for spam detection. Major approaches adopted closer to junk mail filtering encompass "text analysis, white and blacklists of domain names, and community-primarily based techniques". Text assessment of contents of mails is an extensively used method to the spams. Many answers deployable on server and purchaser aspects are available. Naive Bayes is one of the utmost well-known algorithms applied in these procedures. However, rejecting sends essentially dependent on content examination can be a difficult issue in the event of bogus positives. Regularly clients and organizations would not need any legitimate messages to be lost. The boycott approach has been probably the soonest technique pursued for the separating of spams. The technique is to acknowledge all the sends other than those from the area/electronic mail ids. Expressly boycotted. With more up to date areas coming into the classification of spamming space names this technique keeps an eye on no longer work so well. The white list approach is the approach of accepting the mails from the domain names/addresses openly whitelisted and place others in a much less importance queue, that is delivered most effectively after the sender responds to an affirmation request sent through the "junk mail filtering system".

Spam and Ham: According to Wikipedia "the use of electronic messaging systems to send unsolicited bulk messages, especially mass advertisement, malicious links etc." are called as spam. "Unsolicited means that those things which you didn't asked for messages from the sources. So, if you do not know about the sender the mail can be spam. People generally don't realize they just signed in for those mailers when they download any free services, software or while updating the software. "Ham" this term was given by Spam Bayes around 2001 and it is defined as "Emails that are not generally desired and is not considered spam".

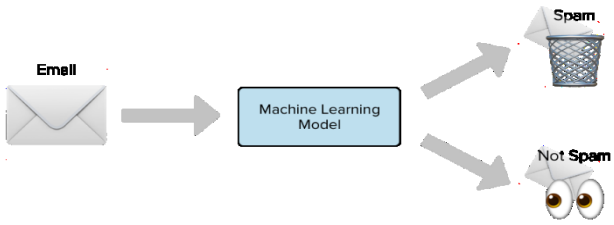


Fig.1. Classification into Spam and non-spam

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:18:34 UTC from IEEE Xplore. Restrictions apply.

Machine learning approaches are more efficient, a set of training data is used, these samples are the set of email which are pre classified. Machine learning approaches have a lot of algorithms that can be used for email filtering. These algorithms include "Naïve Bayes, support vector machines, Neural Networks, K-nearest neighbor, Random Forests etc."

II. LITERATURE REVIEW

There is some related work that apply machine learning methods in email spam detection, A. Karim, S. Azam, B. Shanmugam, K. Kannoorpatti and M. Alazab.[ii] They describe a focused literature survey of Artificial Intelligence Revised (AI) and Machine learning methods for email spam detection. K. Agarwal [3] and T. Kumar. Harisinghaney et al. (2014) [4]and Mohamad & Selamat (2015) [v] have used the "image and textual dataset for the e-mail spam detection with the use of various methods. Harisinghaney et al. (2014) [iv] have used methods of KNN algorithm, Naïve Bayes, and Reverse DBSCAN algorithm with experimentation on dataset. For the text recognition, OCR library" [iii] is employed but this OCR doesn't perform well. Mohamad & Selamat (2015) [v] uses the feature selection hybrid approach of TF-IDF (Term Frequency Inverse Document Frequency) and Rough pure mathematics.

A. Data Set

This model has used email data sets from different online websites like Kaggle, sklearn and some data sets are created by own. A spam email data set from Kaggle is used to train our model and then other email data set is used for getting result "spam.csv" data set contains 5573 lines and 2 columns and other data sets contains 574,1001,956 lines of email data set in text format.

III. METHODOLOGY

A. Data preprocessing:

When the data is considered, always a very large data sets with large no. of rows and columns will be noted. But it is not always the case the data could be in many forms such as Images, Audio and Video files Structured tables etc.
Machine doesn't understand images or video, text data as it is, Machine only understand 1s and 0s.
Steps in Data Preprocessing:
Data cleaning: In this step the work like filling of "missing values", "smoothing of noisy data", "identifying or removing outliers ", and "resolving of inconsistencies is done."
Data Integration: In this step addition of several databases, information files or information set is performed.
Data transformation: Aggregation and normalization is performed to scale to a specific value
Data reduction: This section obtains a summary of the dataset which is very small in size but so far produces the same analytical result

1. Stop words:

"Stop words are the English words that do not add much meaning to a sentence." They can be safely ignored without forgoing the sense of the sentence.
For example if it is tried to search a query like" How to make a veg cheese sandwich", the search engine will try to search the web pages that contains the term "how", "to" ,"make", "a" ,"veg", "cheese" ,"sandwich". The search engine tries to find the web pages that contains the term "how" ,"to", "a" than page containing the recipes of veg cheese sandwich because the terms " how" ,"to", "a" are so commonly used in English language ,If these three words are removed or stopped and actually focuses on retrieving pages that contains the keyword " veg", "cheese", "sandwich" – that would give the result of interest.

2. Tokenization:

"Tokenization is the process of splitting a stream of manuscript into phrase, symbols, words, or any expressive elements named as tokens." The rundown of token further utilized for contribution for additional handling, for example, content mining and parsing. Tokenization is valuable in both semantics (where it is as content division), and as lexical examination in software engineering and building.
It is occasionally hard to define what is intended by the term "word". As tokenization happens at the word level. Frequently a token trusts on modest heuristics, for instance:
Tokens are parted by whitespaces characters, like "line break" or "space", or by "punctuation characters".
Every single neighboring string of alphabetic characters are a piece of one token; similarly, with numbers.
White spaces and punctuations might or might not involve in the resulting lists of tokens.

3. Bag of words

"Bag of Words (BOW) is a method of extracting features from text documents. Further these features can be uses for training machine learning algorithms. Bag of Words creates a vocabulary of all the unique words present in all the document in the Training dataset."

B. CLASSIC CLASSIFIERS

Classification is a form of data analysis that extracts the models describing important data classes. A classifier or a model is constructed for prediction of class labels for example:

"A loan application as risky or safe."

Data classification is a two-step

- learning step (construction of classification model.) and
- a classification step

1. NAÏVE BAYES:

Naïve Bayes classifier was used in 1998 for spam recognition. The Naïve Bayes classifier algorithm is an algorithm which is used for supervised learning. The Bayesian classifier works on the dependent events and works on the probability of the event which is going to occur in the future that can be detected from the same event which occurred previously. Naïve Bayes was made on the Bayes theorem which assumes that features are autonomous of each other. Naïve Bayes classifier technique

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:18:34 UTC from IEEE Xplore. Restrictions apply.

can be used for classifying spam emails as word probability plays main role here. If there is any word which occurs often in spam but not in ham, then that email is spam. Naive Bayes classifier algorithm has become a best technique for email filtering. For this the model is trained using the Naïve Bayes filter very well to work effectively. The Naive Bayes always calculates the probability of each class and the class having the maximum probability is then chosen as an output. Naïve Bayes always provide an accurate result. It is used in many fields like spam filtering.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1)$$

$$P(B) = \sum_{y} P(B|A)P(A) \quad (2)$$

2. *SUPPORT VECTOR MACHINE*

"The Support Vector Machine (SVM) is a popular Supervised Learning algorithm, the Support Vector model is used for classification problems in Machine Learning techniques. "The Support Vector Machines totally founded on the idea of Decision points. The Main resolution of Support Vector Machine algorithm is to create the line or decision boundary. The Support Vector Machine algorithm gives hyperplane as a output which classifies new samples. In 2-dimensional space "hyperplane is line dividing a plane into 2 parts where each class is present in one side."

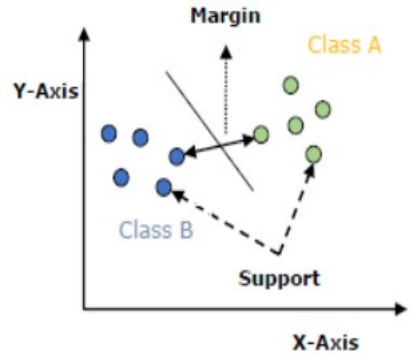


Fig.2 Support Vector Machine

3. *DECISION TREE*

"Decision tree induction is the learning of decision tree from class labeled training tuples". A decision tree is a flow chart like construction, where.

Internal node or non- leaf node= Test on attribute

Branch = shows outcome of the test

Leaf node= holds a class label

Top node is called root node.

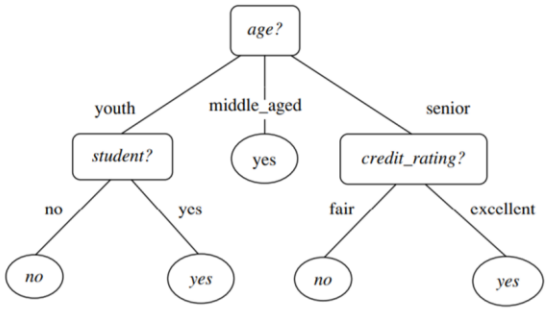


Fig.3. Decision Tree Structure

Decision tree Induction:

The building of "decision tree classifiers" doesn't need "any domain knowledge or parameter setting that is suitable for examining knowledge. "It handles multidimensional information. the learning and classification phases of decision tree induction are simple and fast. Characteristic choice events are utilized to choose the characteristic that top parcel the tuple into particular classes. At the point when choice tree is manufactured a significant number of the branches may result may reflect commotion and anomalies in the preparation information. tree pruning endeavors to recognize and evacuate such branches, with the objective of improving classifier precision on an inconspicuous information.

Entropy using the frequency table of one attribute:

$$E(S) = \sum_{i=1}^{c} - p_i \log_2 p_i \quad (3)$$

Entropy using the frequency table of two attributes:

$$E(T,X) = \sum_{c \in X} P(c)E(c) \quad (4)$$

4. *K- NEAREST NEIGBOUR*

"K-nearest neighbors is a supervised classification algorithm. This algorithm has some data point and data vector that are separated into several classes to predict the classification of new sample point."

K- Nearest neighbor is a LAZY algorithm LAZY algorithm means it tries to only memorize the process it doesn't learn by itself. It doesn't take its own decision by itself.

K- Nearest neighbor algorithm classifies new point based on a similarity measure that can be Euclidian distance.

The Euclidean distance measure Euclidian distance and identifies who are its neighbors.

$$dist((x, y), (a, b)) = \sqrt{(x - a)^2 + (y - b)^2} \quad (5)$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:18:34 UTC from IEEE Xplore. Restrictions apply.

C. ENSEMBLE LEARNING METHODS

"Ensemble methods in machine learning is a method that takes several base model to produce a predictive model in order to decrease. "variance by using bagging bias by using boosting predictions using stacking. Two Types Sequential- here base classifier are created sequentially Parallel- here base classifiers are in parallel.

1. RANDOM FOREST CLASSIFIER

Random forest classifier is an ensemble tree classifier consisting of different types of decision trees that are of different shape and sizes.

The random sampling of the training data when building a tree. A random subgroups of input features when splitting at node in a tree. If you have randomness, the randomization will make look the decision tree less corelated so that generalization error (features of the tree should not look same) of ensemble can be improved.

2. BAGGING

"Bagging classifier is an ensemble classifier that fits base classifiers each on random sub sets of the original data sets and then combined their individual calculations by voting or by averaging) to form a final prediction. "Bagging is a mixture of bootstrapping and aggregating.

Bagging= **B**ootstrap **AGG**regat**ING**

Bootstrapping helps to lessening the variance of the classifier and it also decline the overfitting by just resampling the data from the training data with same cardinality as in original data set. High variance is not good for the model. Bagging is very effective method for limited data, and by just using samples you are able to get estimate by aggregating the scores.

3. BOOSTING AND ADABOOST CLASSIFIER

"Boosting is a ensemble method that is use to create a strong classifier using a number of weak classifier. Boosting is complete by creation a model from a training data sets, then create another model that will precise the faults of the first model." [8] In Boosting Model are added till the training set is predicted properly.

AdaBoost= **Ada**ptive **Boost**ing

AdaBoost is a first fruitful boosting algorithm that was settled for binary classification. The boosting is understood by using AdaBoost.

IV. ALGORITHMS

1.1. Insert the dataset or file for training or testing.
1.2. Check the dataset for supported encoding.
  1.2.1. If one of the supported encodings, then go to step **1.4.**
  1.2.2. If not one of the supported encoding, then go to step **1.3.**
1.3. Change the encoding of the inserted file into one of the supported encodings. Then try again for reading.
1.4. Select whether you want to "Train", "Test' or "Compare" the models using the dataset.
  1.4.1. If "Train" is selected, then go to step **1.5.**
  1.4.2. If "Test" is selected, then go to step **1.6.**
  1.4.3. If "Compare" is selected, then go to step **1.7.**

1.5. "Train" selected:
  1.5.1. Select which classifier to train using the inserted dataset.
  1.5.2. Check for duplicates and NAN values.
  1.5.3. Find the values from Hyperparameter Tuning.
  1.5.4. Process the text for feature transform.
  1.5.5. Train the model
  1.5.6. Save the model and features. Show the results.
  1.5.7. Select which classifier to test using the inserted dataset.
  1.5.8. Check for duplicates and NAN values.
  1.5.9. Load the model and features saved in the training phase of the model.
  1.5.10. Using the loaded values for testing the dataset.
  1.5.11. Show the results

1.6. "Compare" selected:
  1.6.1. Compare all the classifiers using the inserted dataset.
  1.6.2. Show the results of the classifiers.

A. Implementation

Visual studio code platform is used to implement the model and, in this module, a dataset from "Kaggle" website is used as a training dataset. The inserted dataset is first checked for duplicates and null values for better performance of the machine. Then, the dataset is split into 2 sub-datasets; say "train dataset" and "test dataset" in the proportion of 70:30. Then the "train" and "test" dataset is then passed as parameters for text-processing. In text-processing, punctuation symbols and words that are in the stop words list are removed and returned as clean words. These clean words are then passed for "Feature Transform". In feature transform, the clean words which are returned from the text-processing are then used for 'fit' and 'transform' to create a vocabulary for the machine. The dataset is also passed for "hyperparameter tuning" to find optimal values for the classifier to use according to the dataset.
After acquiring the values from the "hyperparameter tuning", the machine is fitted using those values with a random state. The state of the trained model and features are saved for future use for testing unseen data.
Using classifiers from module sklearn in python, the machines are trained using the values obtained from above.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:18:34 UTC from IEEE Xplore. Restrictions apply.

B. FlowChart of the model

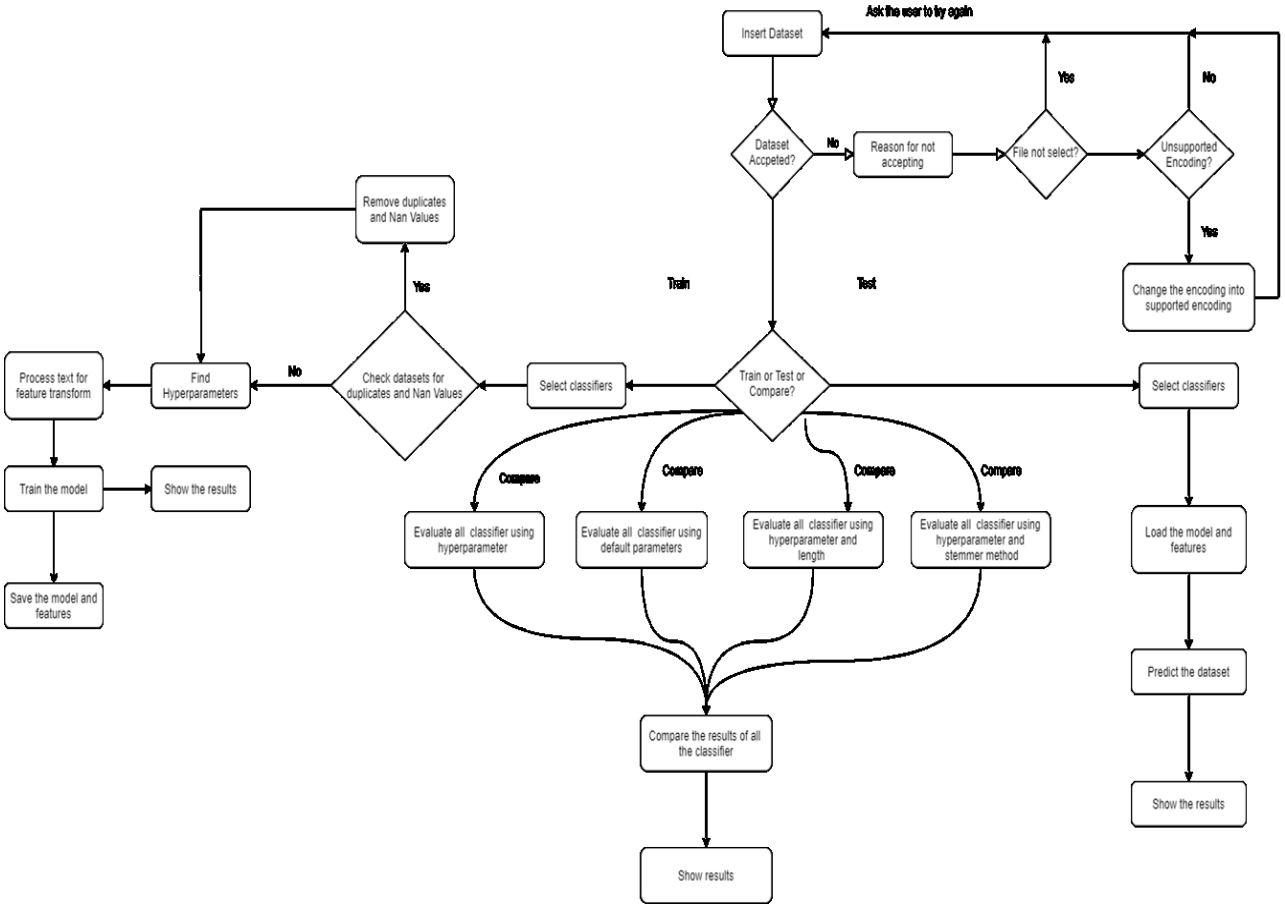


Fig.4. Flow Chart of Model

V. RESULT

Our model has been trained using multiple classifiers to check and compare the results for greater accuracy. Each classifier will give its evaluated results to the user. After all the classifiers return its result to the user; then the user can compare it with other results to see whether the data is “spam” or “ham”. Each classifier result will be shown in graphs and tables for better understanding. The dataset is obtained from “Kaggle” website for training. The name of the dataset used is “spam.csv”. To test the trained machine, a different CSV file is developed with unseen data i.e. data which is not used for the training of the machine; named “emails.csv”. After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:18:34 UTC from IEEE Xplore. Restrictions apply.

TABLE I. COMPARISION TABLE

	Classifiers	*Score 1*	*Score 2*	*Score 3*	*Score 4*
1	Support Vector Classifier	0.81	0.92	0.95	0.92
2	K-Nearest Neighbour	0.92	0.88	0.87	0.88
3	Naïve Bayes	0.87	0.98	0.98	0.98
4	Decision Tree	0.94	0.95	0.93	0.95
5	Random Forest	0.90	0.92	0.92	0.92
6	AdaBoost Classifier	0.95	0.94	0.95	0.94
7	Bagging Classifier	0.94	0.94	0.95	0.94

a. score 1: using default parameters

b. score 2: using hyperparameter tuning

c. score 3: using stemmer and hyperparameter tuning

d. score 4: using length, stemmer and hyperparameter tuning

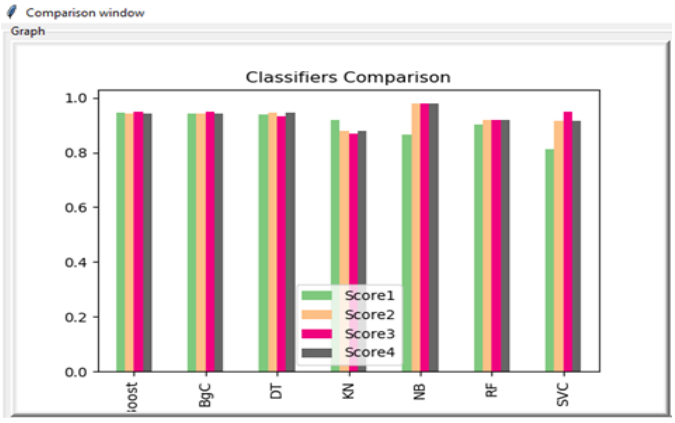


Fig.5 Comparison of all algorithms

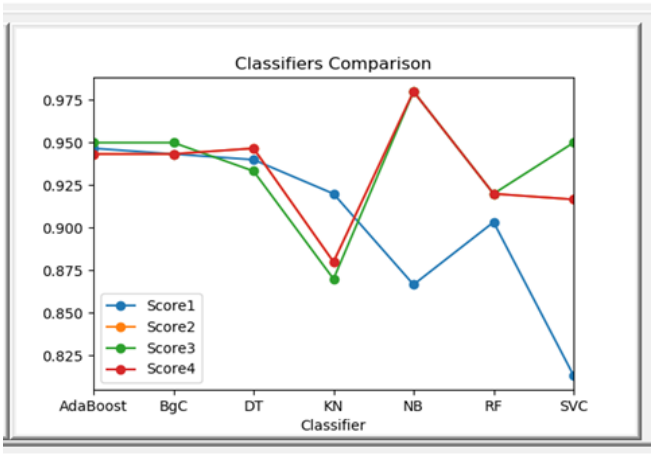


Fig.6. Comparison Graph

VI. CONCLUSION

With this result, it can be concluded that the Multinomial Naïve Bayes gives the best outcome but has limitation due to class-conditional independence which makes the machine to misclassify some tuples. Ensemble methods on the other hand proven to be useful as they using multiple classifiers for class prediction. Nowadays, lots of emails are sent and received and it is difficult as our project is only able to test emails using a limited amount of corpus. Our project, thus spam detection is proficient of filtering mails giving to the content of the email and not according to the domain names or any other criteria. Therefore, at this it is an only limited body of the email.

There is a wide possibility of improvement in our project. The subsequent improvements can be done:

"Filtering of spams can be done on the basis of the trusted and verified domain names."

"The spam email classification is very significant in categorizing e-mails and to distinct e-mails that are spam or non-spam."

"This method can be used by the big body to differentiate decent mails that are only the emails they wish to obtain."

REFERENCES

1. Suryawanshi, Shubhangi & Goswami, Anurag & Patil, Pramod. (2019). Email Spam Detection: An Empirical Comparative Study of Different ML and Ensemble Classifiers. 69-74. 10.1109/IACC48062.2019.8971582.
2. Karim, A., Azam, S., Shanmugam, B., Krishnan, K., & Alazab, M. (2019). A Comprehensive Survey for Intelligent Spam Email Detection. *IEEE Access*, *7*, 168261-168295. [08907831]. https://doi.org/10.1109/ACCESS.2019.2954791
3. K. Agarwal and T. Kumar, "Email Spam Detection Using Integrated Approach of Naïve Bayes and Particle Swarm Optimization," 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2018, pp. 685-690.
4. Harisinghaney, Anirudh, Aman Dixit, Saurabh Gupta, and Anuja Arora. "Text and image-based spam email classification using KNN, Naïve Bayes and Reverse DBSCAN algorithm." In Optimization, Reliabilty, and Information Technology (ICROIT), 2014 International Conference on, pp.153-155. IEEE, 2014
5. Mohamad, Masurah, and Ali Selamat. "An evaluation on the efficiency of hybrid feature selection in spam email classification." In Computer, Communications, and Control Technology (I4CT), 2015 International Conference on, pp. 227-231. IEEE, 2015
6. Shradhanjali, Prof. Toran Verma "E-Mail Spam Detection and Classification Using SVM and Feature Extraction"in International Jouranl Of Advance Reasearch, Ideas and Innovation In Technology,2017 ISSN: 2454-132X Impact factor: 4.295
7. W.A, Awad & S.M, ELseuofi. (2011). Machine Learning Methods for Spam E-Mail Classification. International Journal of Computer Science & Information Technology. 3. 10.5121/ijcsit.2011.3112.
8. A. K. Ameen and B. Kaya, "Spam detection in online social networks by deep learning," 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), Malatya, Turkey, 2018, pp. 1-4.
9. Diren, D.D., Boran, S., Selvi, I.H., & Hatipoglu, T. (2019). Root Cause Detection with an Ensemble Machine Learning Approach in the Multivariate Manufacturing Process.
10. Tasnim Kabir, Abida Sanjana Shemonti, Atif Hasan Rahman. "Notice of Violation of IEEE Publication Principles: Species Identification Using Partial DNA Sequence: A Machine Learning Approach", 2018 IEEE 18th International Conference on Bioinformatics and Bioengineering (BIBE), 2018.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:18:34 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Enhanced Solar PV Power Generation Under PSCs Using Shade Dispersion

Mayank Kumar, *Member, IEEE*

***Abstract*—The solar photovoltaic (PV) array generates the power below the rated value due to the shading intensity and shading pattern losses under partial shading conditions (PSCs). The different levels of solar irradiance are received by the modules of PV array under PSCs; it causes the mismatch between the generated module currents, leading unfortunate impacts such as hot-spot and therefore total PV power generation reduces. Resulting, the efficiency of solar PV array decreases. This article presents the optimal energy harvesting using physical relocation and fixed column position of modules with fixed electrical connection (PRFCPM-FEC) under different shading conditions. The proposed algorithm is presented for $m \times n$ modules under different shading conditions. The locations of modules are changed according to the proposed configuration without removing the electrical connections. A comparative analysis has also been presented between proposed pattern configuration, total cross tied (TCT), and physical relocation of the modules with a fixed electrical connection (PRM-FEC) configurations. The results demonstrated that the proposed array configuration produced the improved performance under PSCs.**

***Index Terms*—Hot-spot, mismatch losses (MLs), partial shading conditions (PSCs), shading patterns, solar array configuration.**

I. Introduction

THE climate change issues, depletion of conventional energy resources, and continuously increasing load demands are creating interest toward clean, natural, and limitless energy resources. In present scenario, solar energy is the best promising resources of clean, limitless, and noiseless energy among all the sustainable energy resources. The major challenges for solar photovoltaic (PV)-based power generation are environmental conditions such as solar irradiance, temperature, nonlinear output characteristics, and low conversion efficiency.

The nonlinear power–voltage (P–V) characteristics of solar PV array under uniform solar irradiance condition contain a unique peak (UP) power point, which is easier to achieve using conventional peak tracking algorithms like perturb and observe, incremental conductance, hill climbing, so on [1]–[3]. Under partial shading conditions (PSCs), the modules of solar PV array receive nonuniform distribution of solar irradiance. The nonuniform distribution of solar irradiance might be caused by shadow of trees/structures, moving clouds, deposition of dust layer on the panels, and so on. Since the generated current mainly depends on the solar irradiance, hence under PSC mismatch of generated current occurs between modules, it limits total generated current from solar PV array and creates local heating regions (hot-spots) [4]. To protect from hot-spot conditions and improve total power generation, the modules are furnished with bypass diodes under PSCs. The average of 41% of total installed solar PV framework is influenced by shading which is responsible for approximately 10% of the total energy loss [5]. The total power harvesting from a PV array is not only depending on the shaded area but it also depends on the shading patterns, location of the shaded modules, array configurations, and so on. The different solar irradiances received by the modules of the PV array also generate multiple peak power points such as local peaks (LPs) and global peak (GP) in $P-V$ characteristics under PSCs. Several GP power point tracking algorithms are presented in the literature [6]–[8]. Therefore, the optimal solar PV power harvesting under various environmental conditions and different shading patterns are major research area for current scenario.

The bypass diodes are used across each module in a PV array configuration to overcome the effect of PSCs at some extent [9]. However, the maximum possible power cannot be achieved using bypass diodes. Several PV array topologies such as series–parallel (SP), bridge link (BL), honey comb (HC), and total cross tied (TCT) have been presented to minimize the partial shading loss (PSL). The TCT array configuration reduces the PSL and it is better array configuration than the other three configurations [9]–[11]. The electronic reconfigurable techniques are used for optimal power generation under PSCs where electrical connection of modules is altered according to the shaded modules but for large PV array installation lots of switches and sensors are required [9], [12]. The limitations of electronic reconfigurable techniques are overcome by physical relocation of the modules (PRM) without disturbing the electrical connection [13] and puzzle pattern arrangement such as Futoshiki [14], and Su Do Ku [15] array configurations. The physical locations of the modules in an array are altered according to the PSCs without altering the electrical connection of the modules. This puzzle pattern arrangement uses TCT topology for $m \times n$ array, where m (number of rows), n (number of columns) are equal for Futoshiki puzzle pattern, and $m = n = 9$ for Su Do Ku puzzle pattern. In [16], an algorithm is also developed for $m \neq n$ and it is reported that power generation is higher than TCT configuration.

Manuscript received June 13, 2020; revised July 27, 2020; accepted August 4, 2020. The review of this article was arranged by Editor B. Hoex.

The author is with the Department of Electrical Engineering, Delhi Technological University, New Delhi 110042, India (e-mail: mayankkumar@dtu.ac.in).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TED.2020.3015171

0018-9383 © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:35:18 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

In this article, an array arrangement algorithm is proposed for optimal energy harvesting using physical relocation and fixed column position of modules with fixed electrical connection (PRFCPM-FEC) under different shading conditions. The TCT configuration is reconfigured with changed physical location of the modules without altering the electrical connection at the time of installation. The further objective of the article is to define physical locations of the modules for $m \times n$; for $m = n$ or $m \neq n$ array configurations. The PRFCPM-FEC configuration is equally valid to define physical locations of the modules for $m \geq n$ or $m \leq n$ solar PV array. This configuration is also designed for m, n = odd, even or both values. Further, the article also presents solar PV array with bypass diodes to reduce local hot-spot conditions. In the PRFCPM-FEC, lots of switches, sensors, and control algorithms are not required (like electronic reconfigurable techniques) for the real-time implementation of the present method in the large-scale solar PV installation.

II. PSL Analysis

The maximum power is generated from a solar PV array under uniform irradiance at standard test condition (UI_STC). The generated power $P_{\text{UI_STC}}$ is approximately equal to the maximum rated power if the system is operated at the UP power point (assume negligible solar power converter losses). The panel generated current (I_{p}) is linearly dependent on solar irradiance and also slightly varies with temperature [17] as given in the following equation, where S and S_{ref} are the absorbed and reference solar irradiation; $\beta = S/S_{\text{ref}}$; I_{p0} is the reference generated current per panel; and T_{c}, $T_{\text{c,ref}}$ are the actual and reference temperature (at STC, $S_{\text{ref}} = 1000$ w/m^2 and $T_{\text{c,ref}} = 25\,^{\circ}$C)

$$I_{\text{p}} = \frac{S}{S_{\text{ref}}}\left\{\text{I}_{\text{p0}} + k_{\text{I}}\left(T_{\text{c}} - T_{\text{c,ref}}\right)\right\} \tag{1a}$$

$$I_{\text{p}} = \frac{S}{S_{\text{ref}}} \times I_{\text{p0}} = \beta \times I_{\text{p0}} \quad \forall T_{\text{c}} = T_{\text{c,ref}}. \tag{1b}$$

Therefore, under PSCs, the generated power per shaded module is lesser than the rated power, and the maximum solar array generated power ($P_{\text{PSC_max}}$) is the sum of individual module power operated at peak power points. The shaded PV module generates lesser module current than the unshaded module, it results in decrease in the overall series connected module current and therefore actual solar array power P_{PSC} is lesser than $P_{\text{PSC_max}}$. The difference between $P_{\text{PSC_max}}$ and P_{PSC} is defined as PSL or more precisely mismatch loss (ML); mathematically, it can be represented as in (2), where N ($= m \times n$) is the number of interconnected solar modules in the array. In addition, multiple peak power points (LPs and GP) appear in $P-V$ characteristics under PSCs which also misled the peak power point tracking algorithm and it is defined as misled losses. All the losses mentioned above are shown in Fig. 1, where P_{ml} and P_{array} are maximum power from lth module and maximum array power under PSCs

$$\text{ML} = \sum_{l=1}^{N} P_{\text{ml}} - P_{\text{array}}. \tag{2}$$

There are several techniques reported in the literature to reduce ML as follows:

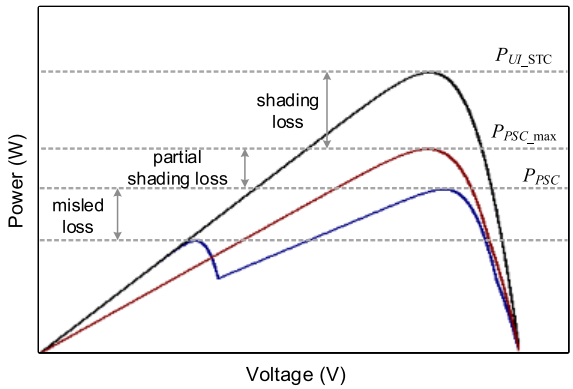


Fig. 1. PSLs (5×5 TCT connected solar PV array configuration and 3×3 SN shading pattern with $\beta = 0.25$).

1) Individual or small group of modules are connected with different maximum power point (MPP) converters to reduce the PSL caused by incoherence between modules. However, additional MPP converters are required for individual modules [18], [19].
2) The independent voltage control of each module is used in multilevel inverters such as diode clamped, capacitor clamped, and cascaded H-bridge to reduce PSL. However, it requires complex control algorithm [20], [21].
3) Reconfiguration of solar PV modules: a) electronic reconfigurable techniques [9], [11], particle swarm optimization (PSO), genetic algorithm (GA)-based reconfigurable techniques [10], so on; however, it requires lots of switches, sensors, and control algorithms and b) fixed electrical connection techniques are used to reduce PSL under PSCs.

The power generated by solar PV array using fixed electrical connection-based technique with TCT configuration under PSC is discussed with respect to voltage and current generations. The current and voltage produced by 5×5 solar array can be written as in the following equations:

$$I_{\text{array}} = \sum_{j=1}^{n} I_{\text{p},ij} \quad \forall\, i = 1, 2, \ldots, (m-1) \tag{3}$$

$$V_{\text{array}} = \sum_{i=1}^{m} V_{\text{p},ij} \quad \forall\, j = 1, 2, \ldots, (n-1). \tag{4}$$

Further, solar PV arrays are considered with TCT configuration and the effect of bypass diodes is discussed under PSCs. Fig. 2(a) shows 5×5 TCT connected solar array under 2×2 shading pattern. Since all the panels from first three rows namely R-1, R-2, R-3 are unshaded, therefore the current generated in these rows are as follows:

$$I_{\text{R-1}} = I_{\text{R-2}} = I_{\text{R-3}} = 5 \times I_{\text{p0}}. \tag{5}$$

The two panels from each fourth and fifth rows (i.e., 41, 42 and 51, 52) are under shading condition. If the shading intensity β is same for all the four panels then the generated current can be written as in the following equation:

$$I_{\text{R-4}} = I_{\text{R-5}} = 2\beta \times I_{\text{p0}} + 3 \times I_{\text{p0}} = (2\beta + 3) \times I_{\text{p0}}. \tag{6}$$

The array contains m ($= 5$) groups of panels that are connected in series where n ($= 5$) panels are connected in parallel

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:35:18 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

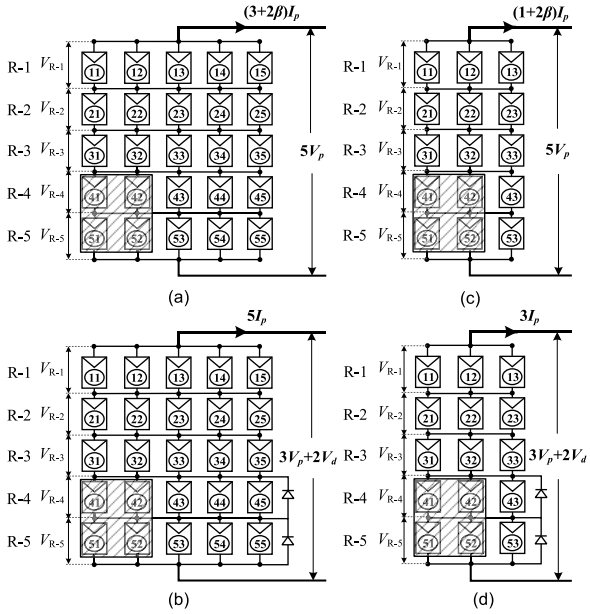


Fig. 2. 5 × 5 TCT connected solar array with 2 × 2 shading pattern (a) without bypass diodes and (b) with bypass diodes. 5 × 3 TCT connected solar array with 2×2 shading pattern (c) without bypass diodes and (d) with bypass diodes.

within the group. Since the array current is limited by series connected shaded groups, therefore the power produced by the array without bypass diodes D_{by} can be written as follows:

$$\mathrm{P}_{\text{array}}\big|_{\text{no}D_{\text{by}}}^{5\times5} = (2\beta+3)I_{\text{p0}}(5V_{\text{p}}) = 15I_{\text{p0}}V_{\text{p}} + 10\beta I_{\text{p0}}V_{\text{p}}. \quad (7)$$

Now, Fig. 2(b) shows the same solar array configuration with bypass diodes D_{by} under same shading pattern. Since bypass diodes are connected in parallel therefore, the current generated by the solar array is the maximum generated current by the solar panel group (i.e., $I_{\text{array}} = 5I_{\text{p0}}$), whereas the array voltage $V_{\text{array}} = 3V_{\text{p}}+2V_{\text{d}}$, where V_{d} is the forward voltage drop across the diode. Since $V_{\text{d}} \ll V_{\text{p}}$, it can be neglected, therefore, the power produced by the array with bypass diodes can be written as follows:

$$P_{\text{array}}\big|_{\text{with}D_{\text{by}}}^{5\times5} = (5I_{\text{p0}})(3V_{\text{p}}) = 15I_{\text{p0}}V_{\text{p}}. \quad (8)$$

Similarly, power can be calculated for 5 × 3 (i.e., $m \neq n$) TCT connected solar array with 2 × 2 shading pattern as shown in Fig. 2(c) and (d) without bypass diodes and with bypass diodes. It can be written as in (9) and (10) with no panels are bypassed and shaded panels are bypassed, respectively

$$P_{\text{array}}\big|_{\text{no}D_{\text{by}}}^{5\times3} = ((1+2\beta)I_{\text{p0}})(5V_{\text{p}}) = 5I_{\text{p0}}V_{\text{p}} + 10\beta I_{\text{p0}}V_{\text{p}} \quad (9)$$

$$P_{\text{array}}\big|_{\text{with}D_{\text{by}}}^{5\times3} = (3I_{\text{p0}})(3V_{\text{p}}) = 9I_{\text{p0}}V_{\text{p}}. \quad (10)$$

Thus, ML for 5 × 5 TCT connected solar array under 2 × 2 short narrow (SN) shading pattern can be calculated from (2) and (7) as follows:

$$\begin{aligned}\text{ML}\big|_{\text{no}D_{\text{by}}}^{5\times5} &= (21+4\beta)I_{p0}V_p - P_{\text{array}}\big|_{\text{no}\ D_{\text{by}}}^{5\times5} \\ &= (21+4\beta)I_{p0}V_p - (15+10\beta)I_{p0}V_p = 6(1-\beta)I_{p0}V_p. \end{aligned} \quad (11a)$$

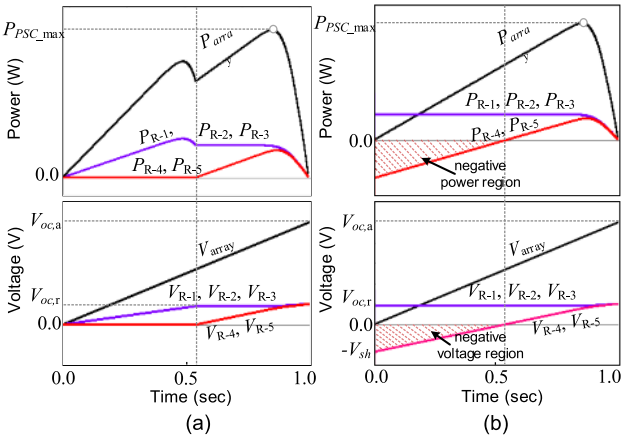


Fig. 3. Power and voltage generated by each row of an array and total power and voltage generated by solar array (5 × 5 TCT solar array with 2 × 2 SN shading pattern) with respect to load variation with time for (a) shaded panels are bypassed and (b) no panel is bypassed.

Similarly, MLs for other configurations can be written using (2), (8)–(10) as in (11b)–(11d), respectively

$$\text{ML}\big|_{\text{with}D_{\text{by}}}^{5\times5} = (6+4\beta)I_{\text{p0}}V_{\text{p}} \quad (11b)$$

$$\text{ML}\big|_{\text{no}D_{\text{by}}}^{5\times3} = 6(1-\beta)I_{\text{p0}}V_{\text{p}} \quad (11c)$$

$$\text{ML}\big|_{\text{with}D_{\text{by}}}^{5\times3} = (2+4\beta)I_{\text{p0}}V_{\text{p}}. \quad (11d)$$

The value of MLs are same for 5 × 5 and 5 × 3 TCT connected solar array with 2 × 2 shaded panel with no panels are bypassed, i.e., without bypass diodes, ML plays dominant role for lower rated solar PV arrays under similar shading conditions. ML decreases when shaded panels are bypassed from 5 × 5 to 5 × 3 TCT connected solar array under 2 × 2 shading pattern, i.e., with bypass diodes, ML decreases for lower rated solar PV arrays under similar shading conditions; also for 5 × 3 configuration the value of ML with D_{by} is lower than ML with no D_{by} for $\beta < 0.4$.

In PSCs with switched-on the bypass diodes, the respective shaded panels are bypassed and therefore, power generated by the shaded panels become zero. Once the unshaded row voltage reaches its open circuit rated voltage of the panel $V_{\text{oc,r}}$, the bypass diode becomes switched-off. Since the bypass diodes are switched-off and therefore voltage produced by the panel increases with respect to the load variation and the generated solar array current flows within the limit of shaded panel current. It can be seen from Fig. 3(a). Additionally, switch-ON state of diodes creates losses due to their ON-state resistance which is considered as negligible.

Fig. 3(b) shows the power and voltage generated by individual rows of solar array and total power and voltage generated by solar array with respect to load variation with time for no panel is bypassed under PSCs. The mismatch of generated current occurs between modules. It creates negative voltage region and therefore produce negative power due to shaded modules, resulting the local heating regions (hot-spots), thus reducing the generation and increasing the total thermal loss. This serious problem deteriorates the solar cell characteristics, protecting shield life and it may damage the complete solar PV module. Therefore, to protect from hot-spot conditions, the modules are furnished with bypass diodes. The maximum

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:35:18 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

negative voltage generated by the row of shaded module V_{sh} can be written as follows:

$$V_{sh} = \frac{(m - m_{sh})V_{oc,r}}{m_{sh}} \quad \forall m_{sh} = 1, 2, \ldots, m \tag{12}$$

where m_{sh} is the number of shaded rows, and $V_{oc,r}$ is the open circuit rated voltage of the panel or the maximum voltage generated by unshaded row of the TCT connected solar array. Therefore, solar PV array consists of solar panels furnished with bypass diodes to reduce local hot-spot conditions; and for the optimal power generation under PSCs or to reduce PSLs (i.e., $P_{PSC_max} - P_{PSC} \approx 0$), PRFCPM-FEC configuration is proposed. It is discussed in Section III.

III. Solar PV Array Configuration for Optimal Power Harvesting

The PRFCPM-FEC methodology is based on fixed electrical connections with relocation of modules for better shading distribution over rows, resulting optimal power can be generated. The proposed algorithm is structured for $m \times n$ solar array, where m = odd or even and n is greater or less than equal to m. Following two representations of solar PV array are used in this article:

$$A = \underbrace{\begin{bmatrix} 11 & 12 & .. & 1n \\ 21 & 22 & .. & 2n \\ .. & .. & .. & .. \\ m1 & m2 & .. & mn \end{bmatrix}}_{\text{Type - I}} = \overbrace{\begin{bmatrix} 1 & 1 & .. & 1 \\ 2 & 2 & .. & 2 \\ .. & .. & .. & .. \\ m & m & .. & m \end{bmatrix}}^{\text{cln1 cln2 .. cln}n}_{\text{Type - II}} \tag{13}$$

$$A_m = \begin{array}{c} \begin{matrix} \text{cln 1} & \text{cln 2} & \text{cln 3} & .. & \text{cln } n \end{matrix} \\ \begin{bmatrix} 1 & k & m & .. & l \\ 2 & k+1 & 1 & .. & l+1 \\ .. & .. & .. & .. & .. \\ k & m & k-1 & .. & l+k-1 \\ k+1 & 1 & k & .. & l+k \\ .. & .. & .. & .. & .. \\ m & k-1 & m-1 & .. & l+m-1 \end{bmatrix} \end{array}, \tag{14}$$

where, $k = \lceil m/2 \rceil \Rightarrow m = 2k - 1$,

$$l = \begin{cases} k - \left\lceil \frac{n}{2} \right\rceil + 1 & \forall\ n = \text{even number} \\ m - \left\lceil \frac{n}{2} \right\rceil + 1 & \forall\ n = \text{odd number.} \end{cases}$$

For m = odd number, the PRFCPM-FEC structure of solar PV array can be achieved by considering following steps:

Step-1: The locations of PV modules for *cln* 1 will be fixed.

Step-2: Divide *cln* 1 into two group of rows, where lower half is started with kth PV module. The position of PV modules for *cln* 2 is structured by rearrangement of lower and upper half of the *cln* 1 with top and bottom position for *cln* 2, respectively. It is presented in (14); where A_m is the modified matrix with m rows, and $m+1 = 1, m+2 = 2, \ldots, m+k = k$ and so on.

Step-3: Similarly, the locations of the PV modules for *cln* 3, 4, $\cdots$,n are obtained; e.g., 5 × 5 solar PV array is obtained and presented in the following equation:

$$\underbrace{\begin{bmatrix} 11 & 32 & 53 & 24 & 45 \\ 21 & 42 & 13 & 34 & 55 \\ 31 & 52 & 23 & 44 & 15 \\ 41 & 12 & 33 & 54 & 25 \\ 51 & 22 & 43 & 14 & 35 \end{bmatrix}}_{\text{Type - I}} \underbrace{\overbrace{\begin{bmatrix} 1 & 3 & 5 & 2 & 4 \\ 2 & 4 & 1 & 3 & 5 \\ 3 & 5 & 2 & 4 & 1 \\ 4 & 1 & 3 & 5 & 2 \\ 5 & 2 & 4 & 1 & 3 \end{bmatrix}}^{\text{cln1 cln2 cln3 cln4 cln5}}}_{\text{Type - II}} \tag{16}$$

Proposed Configurations

For m = even number, two basic matrices $B_{(2,2)}$ and $B_{(3,3)}$ have been used for higher order solar PV array matrices design. It is given in (17). The $(4 \times n)$ and $(6 \times n)$ solar PV arrays can be obtained using $B_{(2,2)}$ and $B_{(3,3)}$ matrices, respectively

$$B_{(2,2)} = \overbrace{\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}}^{\text{cln1 cln2}}; \quad B_{(3,3)} = \overbrace{\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}}^{\text{cln1 cln2 cln3}}. \tag{17}$$

Further, the proposed structure of solar PV array is achieved by considering following steps:

Step-1: The locations of PV modules for *cln* 1 will be fixed.

Step-2: Divide solar array matrix into two groups with respect to rows. The modules position in pth (where p is even) number of column is structured by rearrangement of lower and upper half of the $(P-1)$th column with top and bottom position for pth column, respectively

$$A_4 = \begin{array}{c} \begin{matrix} \text{cln 1} & \text{cln 2} & \text{cln 3} & \text{cln 4} \end{matrix} \\ \begin{bmatrix} 1 & 3_{①} & 2_{②} & 4 \\ 2 & 4_{②} & 1_{①} & 3 \\ 3 & 1_{①} & 4_{②} & 2 \\ 4 & 2_{②} & 3_{①} & 1 \end{bmatrix} \end{array}, \tag{17a}$$

$$A_6 = \begin{array}{c} \begin{matrix} \text{cln 1} & \text{cln 2} & \text{cln 3} & \text{cln 4} & \text{cln 5} & \text{cln 6} \end{matrix} \\ \begin{bmatrix} 1 & 4_{①} & 2_{②} & 5_{②} & 3_{③} & 6 \\ 2 & 5_{②} & 3_{③} & 6_{③} & 1_{①} & 4 \\ 3 & 6_{③} & 1_{①} & 4_{①} & 2_{②} & 5 \\ 4 & 1_{①} & 5_{②} & 2_{②} & 6_{③} & 3 \\ 5 & 2_{②} & 6_{③} & 3_{③} & 4_{①} & 1 \\ 6 & 3_{③} & 4_{①} & 1_{①} & 5_{②} & 2 \end{bmatrix} \end{array}. \tag{17b}$$

Step-3: The modules position in qth (where q is odd) number of columns are structured by rearrangement of lower and upper half of the $(q-1)$th columns with top and bottom position for qth column after the relocation of panel number with respect to basic matrix. e.g., 4 × 4 and 6 × 6 solar PV arrays are obtained by using basic matrix of $B_{(2,2)}$ and $B_{(3,3)}$, and it is presented in (17a) and (17b), respectively.

Step-4: Similarly, $(8 \times n)$, $(10 \times n)$, $(12 \times n)$... solar PV arrays can be obtained by using $B_{(4\times4)}, B_{(5\times5)}, B_{(6\times6)}, \cdots$, basic matrices, respectively.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:35:18 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE I
SPECIFICATIONS OF SOLAR PV MODULE AT STC (IRRADIATION = 1000 W/M^2 AND TEMPERATURE = 25 °C)

Parameters	Values
Solar output power	10 W
Open-circuit voltage V_{OC}	22.50 V
Short-circuit current I_{SC}	0.55 A
Voltage at maximum power V_{MPP}	19.25 V
Current at maximum power I_{MPP}	0.52 A
Number of cells	36 (=9×4) cells
Temperature coefficient of power	-0.43%/ºC
Length, Width, Thickness	255mm, 350mm, 22 mm

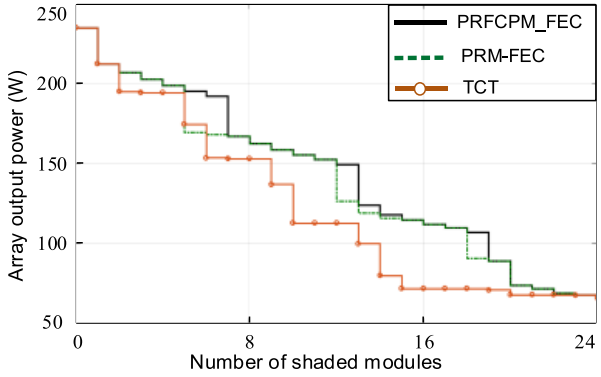


Fig. 4. 6 × 4 array output power with respect to number of shaded modules.

Further, 6 × 4 solar array matrix is designed using (17b) and it is represented in Type-I manner in (17c). In this article two solar PV array sizes [i.e., 5 × 5 in (15), and 6 × 4 in (17c)] have been considered for further analysis. The specification of solar PV module (make Loom Solar Pvt., Ltd, India) is given

$$A_{6\times4} = \begin{bmatrix} 11 & 42 & 23 & 54 \\ 21 & 52 & 33 & 64 \\ 31 & 62 & 13 & 44 \\ 41 & 12 & 53 & 24 \\ 51 & 22 & 63 & 34 \\ 61 & 32 & 43 & 14 \end{bmatrix} \quad (17c)$$

in Table I at STC; where V_{OC} and I_{SC} are open circuit voltage and short circuit current; V_{MPP} and I_{MPP} are voltage and current at MPP, respectively. All the simulation and experimental results are performed/captured with same module parameters.

IV. RESULTS AND DISCUSSION

The power generation comparative analysis of 6 × 4 solar array between proposed algorithm with physical relocation of the modules with a fixed electrical connection (PRM-FEC) [14] and TCT with respect to number of shaded module is shown in Fig. 4. The shading sequence of module followed one by one from left to right and then bottom to top, i.e., 61, 32,···, 51, 22,···, 11,···,54 in proposed configurations as given in (17c); similarly in TCT and PRM-FEC configurations. The result shows that the PRFCPM-FEC relocation of module algorithm generates maximum power.

A small-scale experimental setup was developed to validate the results obtained from mathematical expressions and

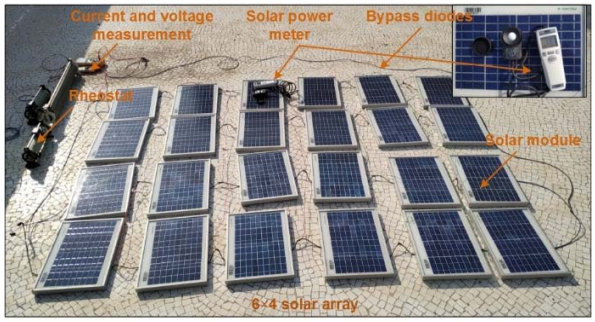


Fig. 5. Experimental setup.

TABLE II
LOCATION OF GP IN TCT AND PROPOSED CONFIGURATION (SN PATTERN)

	TCT Configuration				Proposed Configuration			
	Row current (with respect to bypassed row)		Voltage (V_P)	Power (P_{array})	Row current (with respect to bypassed row)		Voltage (V_P)	Power (P_{array})
(5×5) PV array	$I_{R\text{-}5}$	$2.42\times I_{P0}$	$5\times V_P$	$12.1\times I_{P0}\ V_P$	$I_{R\text{-}5}$	$2.71\times I_{P0}$	$5\times V_P$	$13.6\times I_{P0}\ V_P$
	$I_{R\text{-}4}$	$2.42\times I_{P0}$	-	-	$I_{R\text{-}3}$	$2.71\times I_{P0}$	-	-
	$I_{R\text{-}3}$	$3.0\times I_{P0}$	$3\times V_P$	$9\times I_{P0}\ V_P$	$I_{R\text{-}2}$	$2.71\times I_{P0}$	-	-
	$I_{R\text{-}2}$	$3.0\times I_{P0}$	-	-	$I_{R\text{-}1}$	$2.71\times I_{P0}$	-	-
	$I_{R\text{-}1}$	$3.0\times I_{P0}$	-	-	$I_{R\text{-}4}$	$3.0\times I_{P0}$	$1\times V_P$	$3.0\times I_{P0}\ V_P$
(6×4) PV array	$I_{R\text{-}6}$	$1.9\times I_{P0}$	$6\times V_P$	$11.4\times I_{P0}\ V_P$	$I_{R\text{-}6}$	$2.19\times I_{P0}$	$6\times V_P$	$13.14\times I_{P0}\ V_P$
	$I_{R\text{-}5}$	$1.9\times I_{P0}$	-	-	$I_{R\text{-}5}$	$2.19\times I_{P0}$	-	-
	$I_{R\text{-}4}$	$1.9\times I_{P0}$	-	-	$I_{R\text{-}4}$	$2.19\times I_{P0}$	-	-
	$I_{R\text{-}3}$	$2.48\times I_{P0}$	$3\times V_P$	$7.44\times I_{P0}\ V_P$	$I_{R\text{-}3}$	$2.19\times I_{P0}$	-	-
	$I_{R\text{-}2}$	$2.48\times I_{P0}$	-	-	$I_{R\text{-}2}$	$2.19\times I_{P0}$	-	-
	$I_{R\text{-}1}$	$2.48\times I_{P0}$	-	-	$I_{R\text{-}1}$	$2.19\times I_{P0}$	-	-

PSCAD simulations. Fig. 5 shows 6 × 4 solar PV arrays experimental setup. The setup consists of Loom Solar PV modules of same rating as given in Table I, rheostat, multimeters, solar power meter (make MECO instruments Pvt. Ltd., India), and transparent sheets to create partial shading. The level of solar irradiation is calculated by solar power meter at the time of the experiment and it is verified using (1b) and the parameters mentioned in Table I.

A. Analytical, Simulation, and Experimental Results for Optimal Power Generation

1) Case 1: SN: In SN shading pattern, 5 × 5 and 6 × 4 solar PV arrays are received two different types of solar irradiations 600 w/m^2, 310 w/m^2, and 620 w/m^2, 330 w/m^2, respectively. The shade dispersion for both the solar arrays with respect to PRFCPM-FEC configuration is performed and it is shown in Fig. 6(a) and (b), respectively. The $P-V$ characteristics of solar PV array are plotted using simulation and experimental results for TCT and PRFCPM-FEC with 5 × 5 and 6 × 4 solar array in Fig. 6(c) and (d), respectively.

The analytical solutions (current across each row, voltage and maximum power) for TCT and PRFCPM-FEC are calculated and are presented in Table II. It can be seen that GP power generated from TCT and PRFCPM-FEC configurations for two different solar PV arrays are 12.1 $I_{p0}V_p$, 11.4 $I_{p0}V_p$ and 13.6 $I_{p0}V_p$, 13.14 $I_{p0}V_p$, respectively. In SN shading pattern, the power enhancement for 5 × 5 and 6 × 4 solar PV arrays calculated theoretically from PRFCPM-FEC to TCT configurations are 12.39% and 15.26%, respectively.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:35:18 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

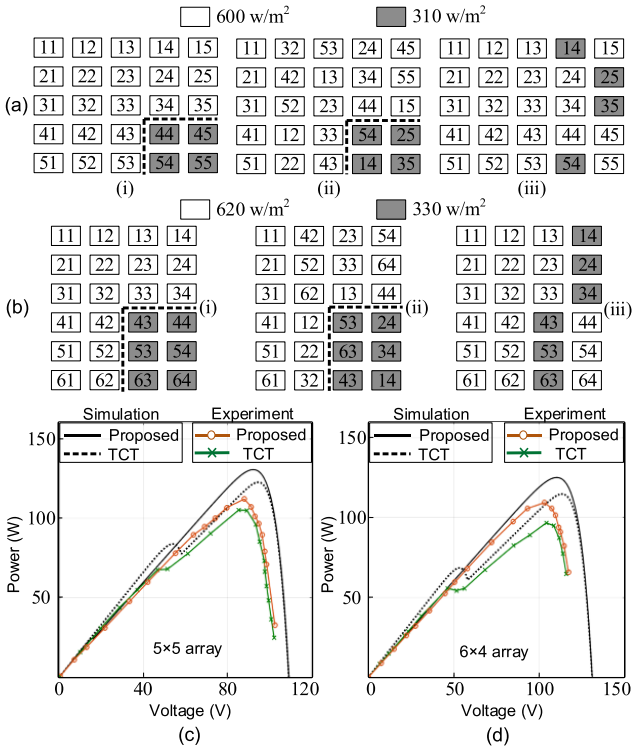


Fig. 6. SN shading pattern with (a) 5 × 5 solar array and (b) 6 × 4 solar array under (i) TCT configuration, (ii) proposed configuration, (iii) shade dispersion with proposed configuration, and (c) and (d) *P–V* characteristics for TCT and proposed configuration with 5 × 5 and 6 × 4 solar array, respectively.

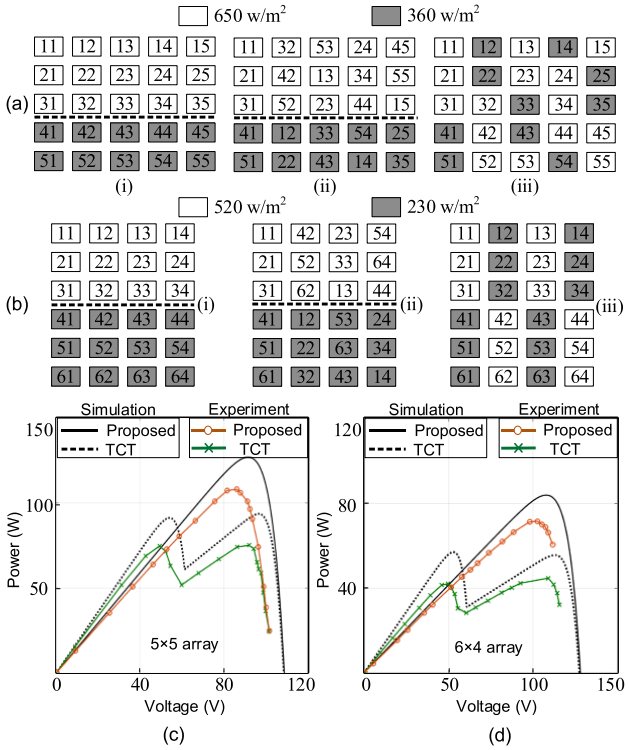


Fig. 7. SW shading pattern with (a) 5 × 5 solar array and (b) 6 × 4 solar array under (i) TCT configuration, (ii) proposed configuration, (iii) shade dispersion with proposed configuration, and (c) and (d) $P-V$ characteristics for TCT and proposed configuration with 5 × 5 and 6 × 4 solar array, respectively.

TABLE III
LOCATION OF GP IN TCT AND PROPOSED CONFIGURATION (SW PATTERN)

	TCT Configuration				Proposed Configuration			
	Row current (with respect to bypassed row)		Voltage (V_P)	Power (P_{array})	Row current (with respect to bypassed row)		Voltage (V_P)	Power (P_{array})
(5×5) PV array	$I_{R\text{-}5}$	$1.8\times I_{P0}$	$5\times V_P$	$9\times I_{P0}\ V_P$	$I_{R\text{-}5}$	$2.67\times I_{P0}$	$5\times V_P$	$13.35\times I_{P0}\ V_P$
	$I_{R\text{-}4}$	$1.8\times I_{P0}$	-	-	$I_{R\text{-}4}$	$2.67\times I_{P0}$	-	-
	$I_{R\text{-}3}$	$3.25\times I_{P0}$	$3\times V_P$	$9.75\times I_{P0}\ V_P$	$I_{R\text{-}3}$	$2.67\times I_{P0}$	-	-
	$I_{R\text{-}2}$	$3.25\times I_{P0}$	-	-	$I_{R\text{-}2}$	$2.67\times I_{P0}$	-	-
	$I_{R\text{-}1}$	$3.25\times I_{P0}$	-	-	$I_{R\text{-}1}$	$2.67\times I_{P0}$	-	-
(6×4) PV array	$I_{R\text{-}6}$	$0.92\times I_{P0}$	$6\times V_P$	$5.52\times I_{P0}\ V_P$	$I_{R\text{-}6}$	$1.5\times I_{P0}$	$6\times V_P$	$9\times I_{P0}\ V_P$
	$I_{R\text{-}5}$	$0.92\times I_{P0}$	-	-	$I_{R\text{-}5}$	$1.5\times I_{P0}$	-	-
	$I_{R\text{-}4}$	$0.92\times I_{P0}$	-	-	$I_{R\text{-}4}$	$1.5\times I_{P0}$	-	-
	$I_{R\text{-}3}$	$2.08\times I_{P0}$	$3\times V_P$	$6.24\times I_{P0}\ V_P$	$I_{R\text{-}3}$	$1.5\times I_{P0}$	-	-
	$I_{R\text{-}2}$	$2.08\times I_{P0}$	-	-	$I_{R\text{-}2}$	$1.5\times I_{P0}$	-	-
	$I_{R\text{-}1}$	$2.08\times I_{P0}$	-	-	$I_{R\text{-}1}$	$1.5\times I_{P0}$	-	-

TABLE IV
LOCATION OF GP IN TCT AND PROPOSED CONFIGURATION (LN PATTERN)

	TCT Configuration				Proposed Configuration			
	Row current (with respect to bypassed row)		Voltage (V_P)	Power (P_{array})	Row current (with respect to bypassed row)		Voltage (V_P)	Power (P_{array})
(5×5) PV array	$I_{R\text{-}5}$	$2.52\times I_{P0}$	$5\times V_P$	$12.6\times I_{P0}\ V_P$	$I_{R\text{-}5}$	$2.52\times I_{P0}$	$5\times V_P$	$12.6\times I_{P0}\ V_P$
	$I_{R\text{-}4}$	$2.52\times I_{P0}$	-	-	$I_{R\text{-}3}$	$2.52\times I_{P0}$	-	-
	$I_{R\text{-}3}$	$2.52\times I_{P0}$	-	-	$I_{R\text{-}1}$	$2.52\times I_{P0}$	-	-
	$I_{R\text{-}2}$	$2.52\times I_{P0}$	-	-	$I_{R\text{-}4}$	$2.81\times I_{P0}$	$2\times V_P$	$5.62\times I_{P0}\ V_P$
	$I_{R\text{-}1}$	$3.1\times I_{P0}$	$1\times V_P$	$3.1\times I_{P0}\ V_P$	$I_{R\text{-}2}$	$2.81\times I_{P0}$	-	-
(6×4) PV array	$I_{R\text{-}6}$	$1.62\times I_{P0}$	$6\times V_P$	$9.72\times I_{P0}\ V_P$	$I_{R\text{-}6}$	$1.62\times I_{P0}$	$6\times V_P$	$9.72\times I_{P0}\ V_P$
	$I_{R\text{-}5}$	$1.62\times I_{P0}$	-	-	$I_{R\text{-}4}$	$1.62\times I_{P0}$	-	-
	$I_{R\text{-}4}$	$1.62\times I_{P0}$	-	-	$I_{R\text{-}3}$	$1.62\times I_{P0}$	-	-
	$I_{R\text{-}3}$	$1.62\times I_{P0}$	-	-	$I_{R\text{-}1}$	$1.62\times I_{P0}$	-	-
	$I_{R\text{-}2}$	$1.62\times I_{P0}$	-	-	$I_{R\text{-}5}$	$1.91\times I_{P0}$	$2\times V_P$	$3.82\times I_{P0}\ V_P$
	$I_{R\text{-}1}$	$2.2\times I_{P0}$	$1\times V_P$	$2.2\times I_{P0}\ V_P$	$I_{R\text{-}2}$	$1.91\times I_{P0}$	-	-

2) Case 2: Short Wide (SW): In SW shading pattern, 5 × 5 and 6 × 4 solar PV arrays are received two different types of solar irradiations 650 w/m^2, 360 w/m^2 and 520 w/m^2, 230 w/m^2, respectively. The shade dispersion for both the solar arrays with respect to PRFCPM-FEC configuration is performed and it is shown in Fig. 7(a) and (b), respectively. The $P-V$ characteristics of solar PV array are plotted using simulation and experimental results for TCT and PRFCPM-FEC with 5 × 5 and 6 × 4 solar array in Fig. 7(c) and (d), respectively.

The analytical solutions (current across each row, voltage, and maximum power) for TCT and PRFCPM-FEC are calculated and are presented in Table III. It can be seen that GP power generated from TCT and PRFCPM-FEC configurations for two different solar PV arrays are 9.75 $I_{p0}V_p$, 7.8 $I_{p0}V_p$ and 13.35 $I_{p0}V_p$, 9 $I_{p0}V_p$, respectively. In SW shading pattern, the power enhancement for 5 × 5 and 6 × 4 solar PV arrays calculated theoretically from PRFCPM-FEC to TCT configurations are 36.92% and 15.38%, respectively.

3) Case 3: Long Narrow (LN): In LN shading pattern, 5 × 5 and 6 × 4 solar PV arrays are received two different types of solar irradiations 620 w/m^2, 330 w/m^2 and 550 w/m^2, 260 w/m^2, respectively. The shade dispersion for both the solar arrays with respect to PRFCPM-FEC

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:35:18 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

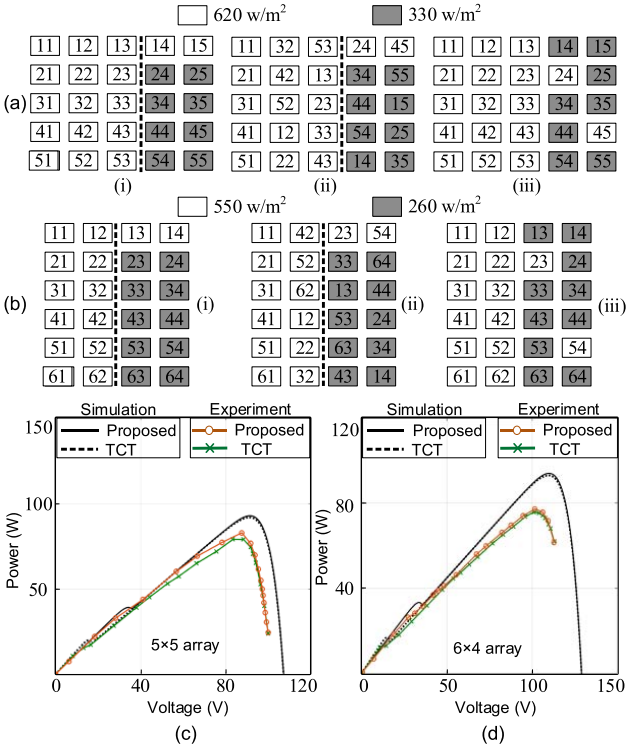


Fig. 8. LN shading pattern with (a) 5 × 5 solar array and (b) 6 × 4 solar array under (i) TCT configuration, (ii) proposed configuration, (iii) shade dispersion with proposed configuration, and (c) and (d) $P-V$ characteristics for TCT and proposed configuration with 5 × 5 and 6 × 4 solar array, respectively.

TABLE V
LOCATION OF GP IN TCT AND PROPOSED CONFIGURATION (LW PATTERN)

	TCT Configuration				Proposed Configuration			
	Row current (with respect to bypassed row)		Voltage (V_P)	Power (P_{array})	Row current (with respect to bypassed row)		Voltage (V_P)	Power (P_{array})
(5×5) PV array	$I_{R\text{-}5}$	$1.8\times I_{P0}$	$5\times V_P$	$9\times I_{P0}\ V_P$	$I_{R\text{-}4}$	$2.09\times I_{P0}$	$5\times V_P$	$10.45\times I_{P0}\ V_P$
	$I_{R\text{-}4}$	$1.8\times I_{P0}$	-	-	$I_{R\text{-}5}$	$2.38\times I_{P0}$	$4\times V_P$	$9.52\times I_{P0}\ V_P$
	$I_{R\text{-}3}$	$2.67\times I_{P0}$	$3\times V_P$	$8.01\times I_{P0}\ V_P$	$I_{R\text{-}3}$	$2.38\times I_{P0}$	-	-
	$I_{R\text{-}2}$	$2.67\times I_{P0}$	-	-	$I_{R\text{-}2}$	$2.38\times I_{P0}$	-	-
	$I_{R\text{-}1}$	$2.67\times I_{P0}$	-	-	$I_{R\text{-}1}$	$2.38\times I_{P0}$	-	-
(6×4) PV array	$I_{R\text{-}6}$	$1.24\times I_{P0}$	$6\times V_P$	$7.44\times I_{P0}\ V_P$	$I_{R\text{-}6}$	$1.53\times I_{P0}$	$6\times V_P$	$9.18\times I_{P0}\ V_P$
	$I_{R\text{-}5}$	$1.24\times I_{P0}$	-	-	$I_{R\text{-}5}$	$1.53\times I_{P0}$	-	-
	$I_{R\text{-}4}$	$1.82\times I_{P0}$	$4\times V_P$	$7.28\times I_{P0}\ V_P$	$I_{R\text{-}3}$	$1.53\times I_{P0}$	-	-
	$I_{R\text{-}3}$	$1.82\times I_{P0}$	-	-	$I_{R\text{-}2}$	$1.53\times I_{P0}$	-	-
	$I_{R\text{-}2}$	$1.82\times I_{P0}$	-	-	$I_{R\text{-}4}$	$1.82\times I_{P0}$	$2\times V_P$	$3.64\times I_{P0}\ V_P$
	$I_{R\text{-}1}$	$1.82\times I_{P0}$	-	-	$I_{R\text{-}1}$	$1.82\times I_{P0}$	-	-

configuration is performed and it is shown in Fig. 8(a) and (b), respectively. The $P-V$ characteristics of solar PV array are plotted using simulation and experimental results for TCT and PRFCPM-FEC with 5 × 5 and 6 × 4 solar array in Fig. 8(c) and (d), respectively.

The analytical solutions (current across each row, voltage, and maximum power) for TCT and PRFCPM-FEC are calculated and are presented in Table IV. It can be seen that GP power generated from TCT and PRFCPM-FEC configurations for two different solar PV arrays are 12.6 $I_{p0}V_p$, 9.72 $I_{p0}V_p$ and 12.6 $I_{p0}V_p$, 9.72 $I_{p0}V_p$, respectively. In LN

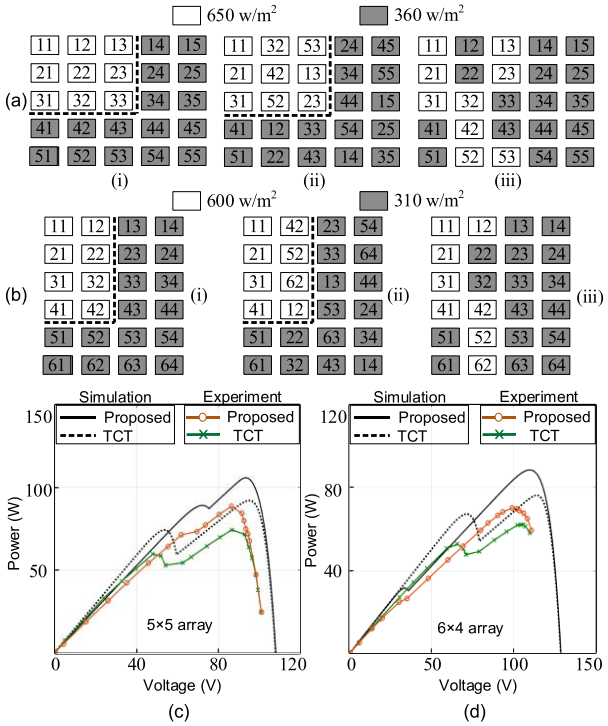


Fig. 9. LW shading pattern with (a) 5 × 5 solar array and (b) 6 × 4 solar array under (i) TCT configuration, (ii) proposed configuration, (iii) shade dispersion with proposed configuration, and (c) and (d) $P-V$ characteristics for TCT and proposed configuration with 5 × 5 and 6 × 4 solar array, respectively.

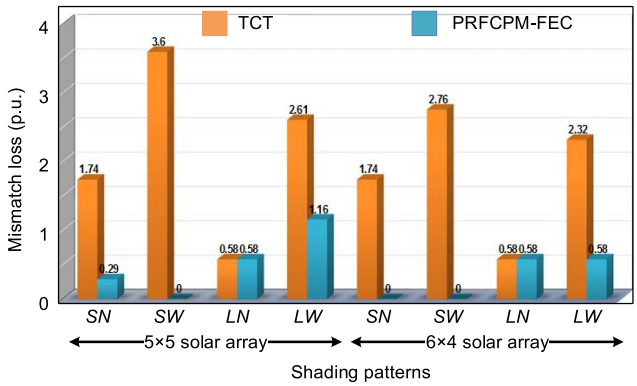


Fig. 10. Comparative results of ML (p.u.) in TCT and PRFCPM-FEC under SN, SW, LN, and LW shading patterns for 5 × 5 and 6 × 4 solar PV array configuration.

shading pattern, the power enhancement from PRFCPM-FEC to TCT configurations is negligible and it is also verified from Fig. 8(c) and (d).

4) Case 4: Long Wide (LW): In LW shading pattern, 5 × 5 and 6 × 4 solar PV arrays are received two different types of solar irradiations 650 w/m^2, 360 w/m^2 and 600 w/m^2, 310 w/m^2, respectively. The shade dispersion for both the solar arrays with respect to PRFCPM-FEC configuration is performed and it is shown in Fig. 9(a) and (b), respectively. The $P-V$ characteristics of solar PV array are plotted using simulation and experimental results for TCT and PRFCPM-FEC with 5 × 5 and 6 × 4 solar array in Fig. 9(c) and (d), respectively.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:35:18 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

The analytical solutions (current across each row, voltage, and maximum power) for TCT and PRFCPM-FEC are calculated and are presented in Table V. It can be seen that GP power generated from TCT and PRFCPM-FEC configurations for two different solar PV arrays are 9 $I_{p0}V_p$, 7.44 $I_{p0}V_p$ and 10.45 $I_{p0}V_p$, 9.18 $I_{p0}V_p$, respectively. In LW shading pattern, the power enhancement for 5 × 5 and 6 × 4 solar PV arrays calculated theoretically from PRFCPM-FEC to TCT configurations are 16.11% and 23.38%, respectively.

B. Calculation of ML

The ML due to change in solar irradiance can be calculated under SN shading condition from Fig. 6 as follows:

$$\mathrm{ML}_{\mathrm{SN}}^{5\times5} = P_{\mathrm{PSC_max}} - P_{\mathrm{PSC}} \tag{18a}$$

$$\mathrm{ML}_{\mathrm{SN}}^{5\times5} = (21\times0.6+4\times0.31) - \{(3\times0.6+2\times0.31)\times5\} = 1.74(\mathrm{p.u.}) \quad \text{TCT configuration} \tag{18b}$$

$$\mathrm{ML}_{\mathrm{SN}}^{5\times5} = (21\times0.6+4\times0.31) - \{(4\times0.6+1\times0.31)\times5\} = 0.29(\mathrm{p.u.}) \quad \text{PRFCPM-FEC configuration} \tag{18c}$$

$$\mathrm{ML}_{\mathrm{SN}}^{6\times4} = (18\times0.62+6\times0.33) - \{2\times(0.62+0.33)\times6\} = 1.74(\mathrm{p.u.}) \quad \text{TCT configuration} \tag{18d}$$

$$\mathrm{ML}_{\mathrm{SN}}^{6\times4} = (18\times0.62+6\times0.33) - \{(3\times0.62+0.33)\times6\} = 0.0(\mathrm{p.u.}) \quad \text{PRFCPM-FEC configuration.} \tag{18e}$$

Similarly, ML under other three shading condition SW, LN, and LW for 5 × 5 and 6 × 4 solar PV arrays are calculated and presented in Fig. 10. It can be observed that ML is minimized and optimal PV power generation can be achieved using proposed algorithm under PSCs. It can also be observed that all the analytical results are very close to simulation results and both these results are in good agreement with experimental results.

V. Conclusion

This article proposed a generalized algorithm for shade dispersion to generate optimal power through solar PV array under PSCs. The PRFCPM-FEC algorithm performs only PRM at the time of PV plant installation. The algorithm is valid for $m \times n$ solar PV array; where $m = n$ or $m \neq n$; $m \geq n$ or $m \leq n$; and $m, n =$ odd, even or both values.

In this technique, the module position is defined only at the time of installation of solar PV array, therefore, the requirement of lots of switches, sensors and other feedback controllers is reduced with respect to electronic reconfigurable techniques. The proposed algorithm generates optimal PV power than the output power generated from TCT, Su Do Ku, Futoshiki puzzle pattern and PRM-FEC configuration under same shading conditions and solar PV structure. The proposed algorithm is applicable for all sizes of solar PV array. The GP power point for 5 × 5 and 6 × 4 solar PV array is calculated theoretically and verified using PSCAD simulations and experimental results.

References

[1] N. Femia, G. Petrone, G. Spagnuolo, and M. Vitelli, "Optimization of perturb and observe maximum power point tracking method," *IEEE Trans. Power Electron.*, vol. 20, no. 4, pp. 963–973, Jul. 2005, doi: 10.1109/TPEL.2005.850975.

[2] F. Liu, S. Duan, F. Liu, B. Liu, and Y. Kang, "A variable step size INC MPPT method for PV systems," *IEEE Trans. Ind. Electron.*, vol. 55, no. 7, pp. 2622–2628, Jul. 2008, doi: 10.1109/TIE.2008.920550.

[3] S. B. Kjær, "Evaluation of the 'hill climbing' and the 'incremental conductance' maximum power point trackers for photovoltaic power systems," *IEEE Trans. Energy Convers.*, vol. 27, no. 4, pp. 922–929, Dec. 2012, doi: 10.1109/TEC.2012.2218816.

[4] D. P. Winston, "Efficient output power enhancement and protection technique for hot spotted solar photovoltaic modules," *IEEE Trans. Device Mater. Rel.*, vol. 19, no. 4, pp. 664–670, Dec. 2019, doi: 10.1109/TDMR.2019.2945194.

[5] E. Koutroulis and F. Blaabjerg, "A new technique for tracking the global maximum power point of PV arrays operating under partial-shading conditions," *IEEE J. Photovolt.*, vol. 2, no. 2, pp. 184–190, Apr. 2012, doi: 10.1109/JPHOTOV.2012.2183578.

[6] A. Ramyar, H. Iman-Eini, and S. Farhangi, "Global maximum power point tracking method for photovoltaic arrays under partial shading conditions," *IEEE Trans. Ind. Electron.*, vol. 64, no. 4, pp. 2855–2864, Apr. 2017, doi: 10.1109/TIE.2016.2632679.

[7] J. S. Goud, R. Kalpana, and B. Singh, "A hybrid global maximum power point tracking technique with fast convergence speed for partial-shaded PV systems," *IEEE Trans. Ind. Appl.*, vol. 54, no. 5, pp. 5367–5376, Sep. 2018, doi: 10.1109/TIA.2018.2845415.

[8] W. Zhang, G. Zhou, H. Ni, and Y. Sun, "A modified hybrid maximum power point tracking method for photovoltaic arrays under partially shading condition," *IEEE Access*, vol. 7, pp. 160091–160100, 2019, doi: 10.1109/ACCESS.2019.2950375.

[9] M. Z. Shams El-Dein, M. Kazerani, and M. M. A. Salama, "Optimal photovoltaic array reconfiguration to reduce partial shading losses," *IEEE Trans. Sustain. Energy*, vol. 4, no. 1, pp. 145–153, Jan. 2013, doi: 10.1109/TSTE.2012.2208128.

[10] T. S. Babu, J. P. Ram, T. Dragicevic, M. Miyatake, F. Blaabjerg, and N. Rajasekar, "Particle swarm optimization based solar PV array reconfiguration of the maximum power extraction under partial shading conditions," *IEEE Trans. Sustain. Energy*, vol. 9, no. 1, pp. 74–85, Jan. 2018, doi: 10.1109/TSTE.2017.2714905.

[11] N. D. Kaushika and N. K. Gautam, "Energy yield simulations of interconnected solar PV arrays," *IEEE Trans. Energy Convers.*, vol. 18, no. 1, pp. 127–134, Mar. 2003, doi: 10.1109/TEC.2002.805204.

[12] J. P. Storey, P. R. Wilson, and D. Bagnall, "Improved optimization strategy for irradiance equalization in dynamic photovoltaic arrays," *IEEE Trans. Power Electron.*, vol. 28, no. 6, pp. 2946–2956, Jun. 2013, doi: 10.1109/TPEL.2012.2221481.

[13] P. S. Rao, G. S. Ilango, and C. Nagamani, "Maximum power from PV arrays using a fixed configuration under different shading conditions," *IEEE J. Photovolt.*, vol. 4, no. 2, pp. 679–686, Mar. 2014, doi: 10.1109/JPHOTOV.2014.2300239.

[14] H. S. Sahu, S. K. Nayak, and S. Mishra, "Maximizing the power generation of a partially shaded PV array," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 4, no. 2, pp. 626–637, Jun. 2016, doi: 10.1109/JESTPE.2015.2498282.

[15] B. I. Rani, G. S. Ilango, and C. Nagamani, "Enhanced power generation from PV array under partial shading conditions by shade dispersion using Su Do Ku configuration," *IEEE Trans. Sustain. Energy*, vol. 4, no. 3, pp. 594–601, Jul. 2013, doi: 10.1109/TSTE.2012.2230033.

[16] H. S. Sahu and S. K. Nayak, "Extraction of maximum power from a PV array under nonuniform irradiation conditions," *IEEE Trans. Electron Devices*, vol. 63, no. 12, pp. 4825–4831, Dec. 2016, doi: 10.1109/TED.2016.2616580.

[17] J. A. Duffie and W. A. Beckman, *Solar Engineering of Thermal Processes*. Hoboken, NJ, USA: Wiley, 2005.

[18] N. Femia, G. Lisi, G. Petrone, G. Spagnuolo, and M. Vitelli, "Distributed maximum power point tracking of photovoltaic arrays: Novel approach and system analysis," *IEEE Trans. Ind. Electron.*, vol. 55, no. 7, pp. 2610–2621, Jul. 2008, doi: 10.1109/TIE.2008.924035.

[19] E. Roman, R. Alonso, P. Ibanez, S. Elorduizapatarietxe, and D. Goitia, "Intelligent PV module for grid-connected PV systems," *IEEE Trans. Ind. Electron.*, vol. 53, no. 4, pp. 1066–1073, Jun. 2006, doi: 10.1109/TIE.2006.878327.

[20] S. Busquets-Monge, J. Rocabert, P. Rodriguez, S. Alepuz, and J. Bordonau, "Multilevel diode-clamped converter for photovoltaic generators with independent voltage control of each solar array," *IEEE Trans. Ind. Electron.*, vol. 55, no. 7, pp. 2713–2723, Jul. 2008, doi: 10.1109/TIE.2008.924011.

[21] A. I. Bratcu, I. Munteanu, S. Bacha, D. Picault, and B. Raison, "Cascaded DC–DC converter photovoltaic systems: Power optimization issues," *IEEE Trans. Ind. Electron.*, vol. 58, no. 2, pp. 403–411, Feb. 2011, doi: 10.1109/TIE.2010.2043041.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:35:18 UTC from IEEE Xplore. Restrictions apply.

The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/1754-2413.htm

Formal vs informal demand for credit by women in urban India: a comparative study in slums of Delhi and Mumbai

Shruti Malik, Girish Chandra Maheshwari and Archana Singh
Delhi School of Management, Delhi Technological University, Delhi, India

Received 4 June 2019
Revised 16 February 2020
26 June 2020
Accepted 19 July 2020

Abstract

Purpose – Over the period, the role of finance has emerged significant in the socio-economic development of the women. There are two major types of finances, i.e. formal and informal ones. Thus, the purpose of this paper is to investigate first the determinants of the demand for credit and then the demand for these credit sources by women especially in urban slums.

Design/methodology/approach – In this study, a primary survey was conducted with the help of a structured questionnaire in slums of two major urban cities in India, i.e. Delhi and Mumbai. In total, 450 individuals were interviewed in each city.

Findings – This paper presents a range of significant socio-economic factors affecting the demand for credit and source of credit by women borrower in Delhi and Mumbai. Despite, the greater emphasis by the government to increase the formal credit utilization, the informal credit is still preferred.

Practical implications – The outcomes of the study are expectedly useful to various policymakers and banks in encouraging women to opt more for the formal credit. The government can follow the research outcomes to scale up the programmes and schemes targeted for women empowerment in urban slums.

Originality/value – The study is unique of its kind in doing a comparative analysis in slums of two differently located urban cities with large slum population.

Keyword Financial inclusion

Paper type Research paper

Introduction

Access to finance has long been debated in the context of economies with persistently high gender disparities (Duflo, 2012). In India, gender disparity is pervasive due to the existing patriarchal system in many parts. Under the patriarchal system, men have more authority and privilege than women over economic entitlement and decision-making (Rawat, 2014). A strong desire for having sons is also seen in many communities, both religious and cultural segments, as they hold the view that male members contribute more toward family reputation and welfare (Robitaille, 2020). Women, on the other hand, are perceived as liabilities, whose responsibilities are mostly restricted towards taking care of family members and household chores. In the past couple of years, the increasing incidence of dowries, female feticides, domestic violence and rapes have lowered the status of women, especially in the northern parts of India (Nieder *et al.*, 2019; Vicente *et al.*, 2020)[1]. Possibly, the socio-economic pressure of changing aspirations due to migration in urban parts can be

Conflict of interest: On behalf of all authors, the corresponding author states that there is no conflict of interest.

Gender in Management: An International Journal
© Emerald Publishing Limited
1754-2413
DOI 10.1108/GM-06-2019-0087

one of the reasons behind the increase in such crimes against women in the last few years. Further, the failure of the judiciary system in securing timely convictions have also compounded the crime rate (Himabindu *et al.*, 2014). Although recently there has been a marginal improvement in terms of changes in workforce participation by women, especially in urban parts, still social and economic empowerment seems to be skewed towards men (Batra and Reio, 2016).

In comparison to men, Indian women's demand for finance is different owing to the social arrangement of unequal power. Usually, men are entrusted with the productive household assets, on the contrary women lack co-ownership of assets and borrow, primarily, for household provisions (Garikipati, 2008a; Guérin *et al.*, 2013). While the role of formal institutions in the inclusive growth of women has emerged as imperative these days. Evidence suggests that inclusion of women under formal financial institutions are likely to improve the family's welfare, decision-making (Sanyal, 2014), investment in businesses (Dupas and Robinson, 2013), and the education and health of children (Sahay and Cihak, 2018). At the macro level, the financial inclusion of women can stimulate economic growth and development (Swamy, 2014; Sahay and Cihak, 2018). Although the gender gap in formal account ownership has significantly reduced in India from 20% in 2014 to 6% in 2017, the gender gap in formal account usage is still persistent, with 54% of adult women reported to having inactive accounts, as per the Global Findex database 2017 (with no withdrawals or deposits made in the past year) (Demirguc-Kunt *et al.*, 2018). In comparison to male-headed households in India, female-headed households have poor formal financial access and usage (Ghosh and Vinod, 2017). The existence of gender and cultural norms (Johnson, 2004), low income-levels (Aterido *et al.*, 2011), poor financial education (Hung *et al.*, 2012) and lack of physical collateral (Ghosh and Vinod, 2017) are some of the major hindrances in accessing formal credit by women. Indian society is primarily patrilineal, but despite laws of equal inheritance of property among sons and daughters, laws are circumvented in wills and sons are preferred for the inheritance of land by many religious groups (Rao, 2017; Bhalotra *et al.*, 2017). Poor women in general, lack the entitlement to physical assets like land or housing due to which they face greater hardship in fulfilling the collateral requirement for formal credit (Delechat *et al.*, 2018). Consequently, informal credit is the preferred option among female heads of households (Ghosh and Vinod, 2017). Informal credit is easily accessible with little or no collateral and with flexible repayment terms. Sometimes, such loans are referred as "instant loans", used for immediate household consumption by women, when men in households find it socially degrading and a matter of pride to borrow small amounts from informal sources such as neighbourhoods and friends (Garikipati, 2017).

Surprisingly, women are the preferred target group among lenders for several reasons. The loan default is seen to be lower when access to credit sources is limited, especially among women. They are perceived to be more credit-worthy due to their willingness to repay, low delinquency (harassment, loss in social prestige), ability to repay and pessimistic approaches to punishments involved with defaults of credit (Mohiuddin and Al-Jalaly, 1993). Further, men are associated more with moral risks and willful defaults (Kabeer, 2001). Alternatively, women are more likely to repay loan amounts in a timely fashion, through strong social ties, networks and cooperation (Basargekar, 2010; Feigenberg *et al.*, 2014).

Although widely studied, most of the literature on the borrowing choices of women are studied in rural contexts (Garikipati, 2009; Fletschner and Kenney, 2014; Islam *et al.*, 2015). There are, indeed, very few studies that have empirically distinguished the demand for formal and informal credit from the woman's perspective (Pal, 2002; Swain, 2002). To our knowledge, the literature fails to provide adequate evidence in suggesting the impact of socio-economic factors affecting the demand for formal and informal credit by women in the

urban context, wherein we expect the credit requirements are more due to the high cost of living. Thus, considering this backdrop, this study has two important research aspects; first, to study the determinants of the demand for credit by women and secondly, to distinguish the demands for formal and informal credit in urban areas. This paper uses a dataset from the slums of two major urban cities in India, namely, Delhi and Mumbai. Mumbai has the largest share of the slum population of India, followed by Delhi (Bag *et al.*, 2016). This study makes two significant contributions to the existing literature on women and finance: first, it highlights a number of socio-economic factors that impact the demand for formal and informal credit, and second, it outlines the possible reasons for such research outcomes. Despite better access to formal financial institutions in urban areas, this study found informal credit is preferred among women borrowers. The estimation results can help both the government and the banks in understanding the current scenario for better policymaking.

The rest of this paper is organized as follows: the second section discusses the literature; the third section details the methodology used to study the research objectives; the fourth section presents the results, the fifth section discusses the results, the sixth section offers the policy implications and the last section presents the concluding remarks.

Literature review

Status of women in India and the role of finance

Traditionally, in India, poor women are far behind men and lack rights and have lower social status. The prevalence of patriarchal norms in many parts expounds on the current status of women, where men in households dictate how women live from birth to death (Chadda and Sinha, 2013). Typically, women under such norms, comply with male household decisions and, subduing their own thoughts and needs, they perform all household duties, taking care of children and husbands (Razvi and Roth, 2004). As a result, gender inequality is evident in many of the social and economic spheres of Indian society, like discrimination in education (Chadda and Sinha, 2013), health care (Sankaran and Madhav, 2011) and the workforce (Kaushik *et al.*, 2014). Increasing women's economic productivity is likely to improve their value and status in the immediate family and provide for a better allocation of resources in productive activities like education, health care and family planning (Bennett, 1992). Noponen (1992) reported that various credit programs to women not only provide a risk cover during a financial crisis but also can help in uplifting their overall self-esteem, status in the home and community. Holvoet (2005), seeking evidence from two credit programs in Southern India, found that women's bargaining powers and stakes in matters related to credit use increases when credit transfers are channeled directly through women's groups. Swain and Wallentin (2012), drawing evidence from a household survey of Self-Help Group (SHGs) in India, argued that challenging social norms in society through microfinance programs can have a positive impact on women's well-being. In this regard, a good strand of theoretical literature asserted that access to finance is inclined to improve women's social status (Garikipati, 2008a; Datta and Sahu, 2018; Sinha *et al.*, 2019). Considering the greater emphasis on finance, it necessitates that the potential sources of finance to fulfill women's financial needs are also studied.

Sources of finance

Women constantly juggle for various sources of debt to fulfill their consumption needs (Guérin, 2014). They may avail finance from formal and informal financial support. The formal finance is sourced from formal financial institutions such as banks and rural credit societies, wherein the primary prerequisite is the collateral term. In general, women's lack of

ownership of productive assets may cause difficulties in fulfilling the requirements of collateral for formal institutions (Karim, 2011). Fletschner (2009) reported that formal institutions may follow biased lending practices due to their general perceptions of women as smaller and less attractive clients. Noticeably, the gender difference in social norms, behavioural attributes, family rights and responsibilities affect women's ability to access formal finance, though, some women, through spousal intermediation, tend to overcome these barriers (Fletschner and Kenney, 2014). Sometimes, women's restricted spatial mobility due to social restriction and confinement to traditional norms may negatively impact their ability to use formal credit (Guerin, 2011). Thus, women who are mostly backed by someone else in their family, particularly men, are more likely to use the formal credit (Balasubramanian *et al.*, 2019).

Consequently, for many women, the informal credit from moneylenders, suppliers and pawnbrokers are preferred because of the immediate disbursement and proximity to the borrower (Berger, 1989). Informal creditors intermittently proxy collateral with peer pressure (Li *et al.*, 2013) trust and reputation (Postelnicu *et al.*, 2014). Though, in certain emergencies, it was also noticed that women who lack appropriate physical assets may mortgage their belongings, like jewelry and dowries, to avail credit from moneylenders (Goedecke *et al.*, 2018). However, there are some major setbacks with borrowing from informal institutions such as unethical practices of loan recovery through violence, intimidation by illegal third parties (Straub, 2005), social friction (Lee and Persson, 2016) and high-interest rates (Mallick, 2012). Though some studies have reported the negative impacts of formal credit market expansion with informal credit utilization (Burgess and Pande, 2005; Giné, 2011; Yuan and Gao, 2012). Simultaneously, several studies have also witnessed the co-existence of formal and informal credit institutions (Garcia *et al.*, 2015; Wu, 2018; Wang, 2019).

Based on the information obtained from the existing literature, it is found that determinants of credit are influenced highly by the individual characteristics (Iregui-Bohórquez *et al.*, 2018; Parappurathu *et al.*, 2019). There is a need to focus more on studies based on individual characteristics and borrowing behaviour. Thus, the theoretical framework for this study comprises several factors which are likely to influence the demand for formal and informal credit by women. Age has an important role in defining the demand for credit (Yuan and Gao, 2012). According to Rahji and Fakayode (2009), age depicts the ability and maturity of the borrower to use credit. Notably, younger individuals are more inclined to demand loans because of their costly life-cycle events such as marriage and children's education. Further, with the increase in household members the level of consumption increases, which is likely to raise the demand for credit (Malapit, 2012). Household income is linked with the demand for credit and some authors who have reported a decrease in demand for credit with an increase in income (Pastrapa and Apostolopoulos, 2015). Though being employed does not affect the demand for loans for men, but it is expected to reduce the demand for loans by women (Malapit, 2012). As, the increase in household income may decrease the demand for consumption loans which are primarily handled by women in many households (Garikipati, 2008a; Fletschner and Kenney, 2014). The past evidence suggests a significant relationship between the household ownership and demand for credit (Crook, 2001; Nguyen, 2007; Mpuga, 2010). Notably women, irrespective of wealth, are usually deprived of loans if they lack possession of land (Balasubramanian *et al.*, 2019). Household expenditure is another determinant which is likely to affect the demand for credit, particularly in urban cities. Women's demands for loans are expected to vary with different stages in their life cycles. Traditionally, in poor households in India, the demand for credit is expected to be low when they are unmarried and are mostly dependent on their

parents or adult family members, while after marriage demand for finance increases, as women are supposed to take care of her family responsibilities (Garikipati, 2008a). Therefore, it is expected that marital status is an important determinant of demand for credit. Further, the number of years spent in education can also impact financial choices. More, educated a woman the better is her chances to recognize investment opportunities that could, in turn, may raise the demand for production credit (Malapit, 2012). The literature reviewed above suggests various factors affecting the demand for credit, which we take up in the empirical model as discussed in detail in the next section.

Methodology and data

This paper is based on primary datasets which were collected from the residents of slums in two major cities in India, i.e. Delhi and Mumbai. To serve the research objectives, a detailed survey was carried out with the help of a structured questionnaire. The questionnaire is divided into two sections, the first aimed at collecting information on the respondent's socio-economic characteristics such as age, income, employment status and expenditure, while the second part contained questions related to the source of borrowing and the frequency of formal and informal borrowing usage. Face-to-face interviews were conducted by the authors themselves with women age 18 years and above. The selection of the individuals within the slums was random, based on approachability and availability. This study was carried out from May 2017 until July 2017. A pilot survey was also carried out for the pre-testing of the questionnaire. In general, there is lesser-known guidance on sample size estimation for the pilot study. Taking cognizance of various past studies (Johanson and Brooks, 2010; Sim and Lewis, 2012) which states that the sample size of 50 can provide satisfactory results for the preliminary survey; therefore, 50 individuals were randomly selected for pilot testing in each selected city. This research adopted multistage cluster sampling. A probability sampling technique like cluster sampling is useful when groups occur naturally in large geographical areas (Bryman and Bell, 2011). In this study, a multi-stage cluster sampling was found appropriate as in both cities there are naturally occurring clusters. In a multistage cluster sampling, the sample was drawn in many stages, with the help of a probability proportional to size (PPS) technique. The PPS technique is a procedure under which the probability of selected a unit is proportionate to the size of the ultimate unit, i.e. the larger clusters have a high probability of selection in comparison to smaller clusters. Thus, to ensure that an individual in a population has an equal probability of selection regardless of the size of their cluster, sample at each stage of cluster sampling is to be selected as per the size. The sampling frame was obtained from the last census of India in the year 2011 (Census of India, 2011). In Mumbai, four different administrative regions (Island City, Western Suburb, Eastern Suburb and Greater Mumbai) with 23 wards and in Delhi four administrative districts (North, West, East and South) with 11 revenue districts served as the sampling unit for further sample collection. In total six clusters, each in Delhi (North, East, South, West, Central and New Delhi) and Mumbai (Matunga, Andheri (East), Malad, Kandivali, Chembur (East) and Bhandup) were selected for the final sample. The information on locations and the approximate number of slum households in each district in these cities were obtained from a variety of civic bodies available at the state level. As, in the last hierarchical stage of the multistage cluster sampling the same number of units are to be sampled randomly (Benett, 1991); therefore, from the selected clusters in each city, 450 women were randomly interviewed (75 in each cluster).

While the sample size calculation was based on the Cochran (1977) method, according to which the sample size was estimated to be 385 in case of Delhi and Mumbai each. The sample size is further increased to accommodate the sampling error, with an additional 15%,

the sample size of approximately 450 is be obtained. Though the information was collected from 450 individuals each in Delhi and Mumbai, after the processing of data a total of 849 valid responses (401 and 448 in Delhi and Mumbai, respectively) were obtained.

Estimation model

This study sets the estimation model for the research objectives in two steps:

(1) probability of overall demand for credit; and

(2) the probability of choosing between formal and informal credit sources by women.

Predominantly, there are two choices for estimation, i.e. probit and logit models, although, both models yield similar outcomes (Davidson and MacKinnon, 2004). The logistic regression model (or logit model) is particularly useful when the dependent variable is a binary measure (either 0 or 1) and the independent variables are continuous. Moreover, the multinomial logit regression is used when there are more than two classes of dependent variables. For step one, the dependent variables were defined as the respondent who took out the loan in the previous year (value "1") and did not take out a loan (value "0"). With a single dependent variable, the logit model is defined by the expression below:

$$\mathrm{Pi} = \mathrm{f}(\alpha + \beta \mathrm{X}_i) = \frac{1}{1 + \left(\frac{1}{e^{\alpha+\beta x_i}}\right)} \tag{1}$$

e = natural logarithm (2.718);
Pi = Probability of demand for credit (1 for having a loan or 0 for not having a loan); and
Xi = attribute of the ith respondent.

The logit model as represented in equation (1) can be written in terms of the odds ratio. The odd ratio enables the understanding of the interpretations of the coefficients in the logit model. The odd ratio can be represented as in equation (2):

$$\left(\frac{P_i}{1 - P_i}\right) = \left(\frac{1 + e^{Z_i}}{1 + e^{-Z_i}}\right) = e^{Z_i} \tag{2}$$

While Z_i can be defined as per equation (3) (Hosmer *et al.*, 1989):

$$Z_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + + \beta_n X_n + \mu_i \tag{3}$$

β_0 and β_n are the estimated vectors, $\mathrm{X_n}$ represents explanatory variables and μ_i the error term.

To study the first objective, i.e. the probability of taking a loan can be obtained by substituting equation (3) with the explanatory variables (Table 1) as below:

$$\begin{aligned} Z_i(PrCredit) = {} & \beta_0 + \beta_1 Age + \beta_2 Age(Sq) + \beta_2 Education + \beta_3 FamilySize + \beta_4 Income \\ & + \beta_5 Expenditure + \beta_6 HouseholdOwnership + \beta_7 Employment\ \ Status \\ & + \beta_8 MaritalStatus + \mu_i \end{aligned} \tag{4}$$

To study the second research objective, a maximum likelihood technique was used to build an unordered multinomial logit model as in equation (5):

Variable	Description of the variable (along with codes)	Literature
Age	The respondent's age	(Mpuga, 2010; Yuan and Gao, 2012; Pastrapa and Apostolopoulos, 2015)
Age (Sq)	Age squared	
Education	A dummy variable that takes the value 1 if the respondent can read and write (upto 5th standard and above) and 0 otherwise (illiterate)	(Kumar *et al.*, 2010; Yuan and Gao, 2012)
Family size	Number of family members	(Nguyen, 2007; Pastrapa and Apostolopoulos, 2015)
Income	Respondent's household income	(Garikipati, 2008a)
Expenditure	Respondent's household expenditure	(Kaboski and Townsend, 2012)
Household ownership	A dummy variable taking the value 1 if the dwelling is owned, else zero	(Mpuga, 2010)
Employment status	A dummy variable taking the value 1 if the respondent is employed, else zero	(Fidrmuc *et al.*, 2013)
Loan	A dummy variable taking value 1 if the respondent has availed loan in past, else zero	(Pastrapa and Apostolopoulos, 2015)
Loan amount	last loan amount	(Chaudhuri and Cherical, 2012; Islam *et al.*, 2015)
Marital status	A dummy variable taking value 1 if the respondent is married, else zero	(Pastrapa and Apostolopoulos, 2015)

Source: Author's compilation

Table 1. Details of the variables used in the empirical analysis

$$Prob[Y = j] = \frac{e^{\beta_j X_i}}{1 + \sum_{k=0}^{3} e^{\beta_j X_i}} \qquad (5)\ (j = 0,1,2)$$

where Y is the dependent variable, which represents the probability of demand for choosing a credit source (formal or informal). The dependent variable (Y) is classified under three groups i.e. "0" when the ith individual did not take out a loan, "1" when the ith individual took out a formal loan (banks, formal microfinance institutions, etc.) and "2" when the ith individual has an informal loan (borrowing from friends/family members, moneylender, etc.). All the explanatory variables selected are listed in Table 1.

Results and analysis

Descriptive statistics

The sample distributions in the two cities are depicted in Table 2.

A large number of respondents are reported in the age group of 29–38 years in Delhi and Mumbai. In Mumbai, the maximum monthly household earnings were between INR 9001 and 12000, while in Delhi, the maximum income ranged between INR 6001 and 9000. The average household size was four in Delhi and five in Mumbai. While the maximum household expenditure ranged between INR 8001 and 12000 and INR 4001 and 8000 in Mumbai and Delhi, respectively. Respondents resided either in owned or rented houses. In both cities, the respondents lived mostly in rented houses. The number of years spent on schooling varied in Delhi (mostly up to 5th standard) and Mumbai (mostly up to 10th standard). It was found that 71% of female respondents possess bank accounts in Delhi,

City Descriptive	Delhi Frequency (%)	Mumbai Frequency (%)
Age		
18–28	22 (5.5)	22 (4.9)
29–38	253 (63.1)	253 (56.5)
39–48	107 (26.7)	144 (32.1)
49–58	15 (3.7)	23 (5.2)
59 and above	4 (1)	6 (1.3)
Family member		
1–3	33 (8.3)	55 (12.3)
4–6	262 (65.3)	221 (49.3)
7–10	91 (22.6)	124 (27.7)
10 and Above	15 (3.8)	48 (10.7)
Household monthly income (In Indian Rupees)		
1000–3000	61 (15.3)	32 (7.4)
3001–6000	127 (31.6)	56 (13.9)
6001–9000	149 (37.2)	131 (32.6)
9001–12000	39 (9.7)	145 (36.2)
12001 and above	25 (6.2)	84 (20.9)
Household monthly expenditure (In Indian Rupees)		
0–4000	21 (5.2)	7 (1.6)
4001–8000	251 (62.5)	75 (16.7)
8001–12000	89 (22.2)	163 (36.3)
12001–16000	23 (5.7)	107 (23.9)
16000 and above	17 (4.4)	96 (21.5)
Education		
Illiterate	94 (23.5)	121 (27.3)
Upto 5th standard	186 (46.4)	88 (19.6)
Upto 10th standard	96 (23.9)	158 (35.2)
Upto 12th standard	16 (3.9)	60 (13.3)
Graduate and above	9 (2.3)	21 (4.6)
Household ownership		
Rented	247 (61.5)	365 (81.4)
Owned	154 (38.5)	83 (18.5)
Possess bank account		
No	117 (29.2)	55 (12.3)
Yes	284 (70.8)	393 (87.8)
Borrowing source		
Formal	28 (6.9)	35 (7.8)
Informal	70 (17.4)	99 (22.1)
Did not borrow	303 (75.7)	314 (70.1)

Table 2.
Sample descriptive (Delhi and Mumbai)

Note: Total number of observations in case of Delhi and Mumbai are 401 and 448, respectively
Source: Authors estimation

whereas 88% of the sample in Mumbai possessed formal bank accounts. However, the incidence of formal borrowing was lower in both cities. In Mumbai, 22.8 and 7.81% of respondents borrowed from informal and formal credit sources, respectively. In Delhi, 6.98% of individuals borrowed formally and 17.46% informally, although, a large number of women had not borrowed in the previous 12 months.

Empirical estimations

Demand for credit. The estimation results highlight the determinants of credit demand in Delhi and Mumbai as represented in Table 3.

In comparison to Mumbai, age affects the demand for credit in Delhi, with a one-unit change in age the odds of having a loan increase by 2.5 times in Delhi. However, with age (squared) the probability to take out a loan decreases by 3% in Delhi. Theoretical evidence also suggests a positive impact of age on demand for credit (Crook, 2001; Nguyen, 2007; Pastrapa and Apostolopoulos, 2015). However, as per the Modigliani Life-Cycle hypothesis (Ando and Modigliani, 1963; Deaton, 2005), the nature of consumption varies with age, as age progresses people tend to become more risk-averse and are less likely to borrow. This study found education as an important determinant in the context of both Delhi and Mumbai, conversely, education plays an essential role in women making financial decisions. Further, family size exhibits a positive relationship with the demand for credit in Mumbai, with an increase in family size the demand for credit may increase by 1.21 times. This is expected as an increase in family size may raise the household consumptions, which in turn impact the credit needs. Similarly, the household expenditure also increases the odds of having credit in Delhi and Mumbai. Though household income is significant in Delhi and Mumbai, it neither has a positive or negative influence on the odds for borrowing by women. The employment status (being employed) is found to be statistically in both cities, which implies that working women are more likely to demand credit. In support of this, many researchers also confirmed a significant association between various employment options, such as self-employment and wage labourers, with empowerment and improved status of women in the family (Pitt *et al.*, 2006; Banerjee *et al.*, 2015). The marital status (being married) is associated with the probability of credit demand in Mumbai only, the odds of credit demand increases when women are married. Notably, the location variable was found

Loan availed (Yes = 1, No = 0)	Delhi Odd ratio	Mumbai Odd ratio	Combined (Delhi and Mumbai) Odd ratio
Age	2.51** (1.07)	0.88 (0.92)	0.86* (0.08)
Age (Squared)	0.97** (0.00)	1.00 (0.00)	1.00 (0.00)
Education	2.80*** (0.84)	2.83*** (0.72)	3.14*** (0.67)
Family size	1.19 (0.19)	1.21* (0.15)	1.23** (0.13)
Household income	1.00*** (0.00)	1.00*** (0.00)	1.00*** (0.04)
Household expenditure	1.99** (0.00)	2.29** (0.00)	2.18** (0.00)
Household ownership	0.90 (0.66)	0.62 (0.38)	0.54 (0.27)
Employment status	2.81*** (0.83)	1.88*** (0.45)	1.87*** (0.37)
Marital status	1.29 (0.52)	0.42*** (0.11)	0.08*** (0.02)
Location variable			0.45*** (0.09)
Constant	0.32 (0.40)	0.31 (0.60)	0.39 (0.64)

Notes: ***Significant at 1%, **significant at 5%, *significant at 10%, Values in brackets () represents the error
Source: Authors estimation

Table 3. Logistic regression with odd ratios (demand for credit in Delhi and Mumbai)

to be significant, which indicates that there is a significant regional difference between these two cities.

Demand for credit source. The demand for various credit sources (formal and informal) is estimated through the multinomial logit model (Table 4).

The research findings reveal that age and age (squared) are negatively associated with the demand for informal credit in Delhi. Education, on the other hand, is positively linked with the demand for formal credit in both cities. In addition, it is associated with informal loan demand in Delhi at the 1% level of significance. Family size is positively associated with the likelihood of demands for formal credit by women in Mumbai. Household earnings, notably, impact both formal and informal credit choices in Delhi and Mumbai. Household expenditure is an important determinant of informal borrowing in both cities (at the 15% level of significance). Further, household ownership is insignificant in influencing the demand for either formal or informal credit by poor women. Further, the research finds that the employment status (being employed) is positively correlated with informal credit demand in both cities (at the 1% significance level). Also, being employed impacts the probability of formal credit in Delhi. Marital status (being married) is correlated with the probability of formal as well as informal credit demands in Mumbai. The location variable is statistically significant, which implies that there exists a difference in demand for formal and informal credit in both cities.

Discussion

The results of this study reflect how the demand for credit and credit sources varies with regional differences. Poor women residing in slums are mostly migrants from the rural areas in the neighbouring states who moved, along with their spouses, for a better livelihood. As regards the geographical locations of these megacities, Delhi being close to many North-Central states (Uttar Pradesh, Bihar, Uttarakhand, Himachal Pradesh) and North-Western States (Rajasthan, Haryana, Punjab, etc.) have a large number of inhabitants from these states; however, Mumbai, in addition to the vast population of native migrants (from Maharashtra), has attracted a significant number of migrants from far off regions of North-Central and Southern states (Tamil Nadu, Kerala, Andhra Pradesh, Telangana) due to better job opportunities (Bag *et al.*, 2016). It was observed that women in general from villages in the north-central states like Uttar Pradesh, Bihar were more backward, restricted in economic activities in comparison to women from southern parts such as Andhra Pradesh and Tamil Nadu. This regional variation is seen mainly due to differences in marriage and kinship structures. In northern parts, kinship is more patrilineal in nature and women share fewer ties with their father's households after marriage, in contrast, women in southern counterparts tend to be more connected with her natal house even after marriage, with more freedom and less cultural restrictions (Bloom *et al.*, 2001). This is quite evident in this study, as it was noticed that women who hail from slums in the southern states of Mumbai were mostly outspoken, better educated, actively engaged in various occupations to support their families earnings, while in the Delhi slums, as most of the migrants are from north-western states, women in general, were seen reserved, docile and preferred to stay as housemakers. In Mumbai, women were seen less restricted and dependent on their spouse in making their financial choices in comparison to women in Delhi. As a majority of these urban migrants tend to follow the customs and values as being practiced in their native place which is also manifested in their borrowing behaviour.

Besides, this spatial arrangement in slums also has an important role in contextualizing the relationship with various financial choices (Kar, 2018; Harker *et al.*, 2019). The social relationship which emerges with kin and neighbour and the inability to run out from debt

Table 4.
Multinomial logistic regression (demand for credit sources in Delhi and Mumbai)

Place	Delhi		Mumbai		Combined	
Credit source	Formal (1) Coefficient	Informal (2) Coefficient	Formal (1) Coefficient	Informal (2) Coefficient	Formal (1) Coefficient	Informal (2) Coefficient
Age	0.70 (0.66)	−0.96* (0.41)	−0.21 (0.18)	−0.12 (0.11)	−0.21** (0.12)	−0.26* (0.08)
Age (Square)	−0.01 (0.01)	−0.03** (0.02)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00* (0.00)
Education	1.58* (0.50)	0.84* (0.31)	1.75* (0.49)	0.96 (0.27)	0.70* (0.29)	0.09 (0.18)
Family size	0.23 (0.24)	0.11 (0.17)	0.28* (0.20)	0.16 (0.13)	0.17 (0.14)	0.09 (0.09)
Household income	0.00*** (0.00)	0.00* (0.00)	0.00* (0.00)	0.00* (0.00)	0.00** (0.00)	0.00* (0.00)
Household expenditure	0.00 (0.00)	0.00*** (0.01)	0.00 (0.00)	0.00*** (0.01)	0.00 (0.00)	0.00*** (0.01)
Household ownership	−0.13 (1.12)	0.26 (0.84)	−0.91 (0.86)	−0.41 (0.66)	−0.50 (0.65)	0.11 (0.52)
Employment status	0.84** (0.45)	1.10* (0.31)	0.38 (0.40)	0.82* (0.27)	0.14 (0.28)	0.29** (0.18)
Marital status	−0.23 (0.53)	0.63 (0.46)	1.06* (0.41)	0.83* (0.28)	−0.98* (0.30)	−0.72* (0.21)
Location variable					0.11** (0.27)	0.06** (0.17)
Constant	1.62 (0.81)	0.05 (0.16)	−0.99 (3.25)	−1.24 (2.10)	0.84 (2.29)	2.30 (1.47)

Note: ***15%, **5%, *1% Significance level, Values in brackets () represents the error
Source: Authors estimation

obligations in densely packed spaces in slums defines their borrowing behaviour (Kar, 2018). They most probably follow the financial choices as opted by their neighbours. Fletschner and Carter (2008), also asserted that women's financial opportunities are shaped not just by the individual access to finance but also by other women in the community who obtained capital in similar needs. Social networking, plays an important role in accessing information and opportunities, especially in the context of female migrants in urban areas (Banerjee and Govil, 2018). Though there are a range of formal or informal credit sources available in both cities, in most cases woman approached informal credit channels due to lack of land ownership in illegal or unauthorized slums. Besides, the role of patriarchal norms in limiting access to formal credit by women should not be ignored.

This study elucidates a number of socio-economic determinants of the demand for credit and credit sources. In Delhi, factors such as age, education, household expenditure and income have a high bearing on women's demand for credit. In Mumbai, marital status was also found to be significant in addition to the above-mentioned factors (except age). Maharashtrian society is, in general, opposed to the concept of dowry as compared to the northern counterparts in the country [2] [3]. Therefore, after marriage, demand for loans by women in Mumbai is probably seen as higher, as sometimes, they do not prefer or get dowries to get their houses. Additionally, they like to live in a house separate from their parents due to limited space in Mumbai. Further, when demand for credit sources are examined, this study noticed many striking similarities in Delhi and Mumbai. Education is one such factor that bears a positive relationship with demand for formal credit in these metro cities Education enables women to make more prudent financial choices and better utilization of formal institutional services (Ghosh and Vinod, 2017). In contrast, many authors who reported household ownership as an important determinant for the demand for formal credit (Crook, 2001; Nguyen, 2007; Pastrapa and Apostolopoulos, 2015), this paper found household ownership as statistically insignificant. Possibly, due to the patrilineal system in India, a large proportion of women are seen excluded from property rights. Consequently, the demand for formal credit decreases as they are unable to provide the required bank collateral (Koride and Gurtoo, 2019). Moreover, it can be seen that women living in slums are migrants who look for a roof to sleep under, rather than a house to own, due to poor affordability and their preference to build houses in their home states. Principally, this difference in the demand for credit and sources of credit in Delhi and Mumbai can also be explained through variations in culture, castes and religious beliefs in slum ecosystems (Bag *et al.*, 2016). For instance, Ruthven (2002) noticed that mixed OBC (other backward classes) from western Uttar Pradesh and Muslims were least served in the slum environments of west Delhi squatter settlements [4].

Policy implications

The estimation results may facilitate various stakeholders in governments and banks for better formulation of plans to reduce institutional frailty in fulfilling women's credit demands in urban areas. Bankers need to integrate the identified socio-economic characteristics in their financial products and services. Formal finance is considered more advantageous than informal ones due to certainty, fixed-interest rates and lower risk. However, this study found less than one-third availed institutional credit in both cities. Thus, banks need to consistently innovate their financial products and should work to eliminate the need for physical assets as collateral. Besides, this the government can use these research findings to scale-up programs and schemes targeted for female empowerment. There should be more emphasis on financial literacy programs for women to promote the formal credit usage. Further, the government and banks need to nurture a

conducive environment to promote microenterprises through self-help groups (SHGs). Women located in close-knit spaces in urban slums can benefit to a large extent as it is easier for them to interact and form SHGs.

Future researchers are prompted to investigate the possible reasons for the research outcomes. It is important to highlight that some of the variables under investigation were found to be significant in one geographical setting while not so significant in another area. For instance, education which was found to be a positive indicator of demand for credit in Delhi and was found insignificant in case of Mumbai, possibly due to lower educational attainment. Thus, the socio-economic factors presented should ideally be aligned with geographical idiosyncrasies.

Conclusion

This paper has twin objectives, first, to determine the socio-economic factors influencing the demand for credit and secondly to distinguish between the demand for formal and informal credit by women in the slums of two economically significant cities in India, i.e. Delhi and Mumbai. Factors such as education, employment and household earnings and expenditures were found significant in affecting the demand for credit in both cities. Further, when demand for formal and informal credit was studied, education and household income emerged as important determinants of the demand for formal credit and, on the other hand, being employed, household income and expenditure are the factors which impact the demand for informal credit in the context of Delhi and Mumbai. Although a large proportion of women were having bank account in both cities, still formal credit usage was seen poor. Informal credits were mostly preferred. Thus, the government and banks need to adopt different strategies in both cities to increase demand for formal credit. For instance, education impacts the demand for formal credit in both cities; therefore, financial institutions should design financial services in local languages with more user-friendly interfaces to ease access by poor uneducated women in slums. Financial institutions should come up with more age, income and occupation-specific credit schemes. Though bank-led SHGs models are proven successful for women in many rural areas, especially when loan amounts are small, and collateral is inadequate. More such models need to be replicated by banks to promote entrepreneurship among poor women struggling each day for income in urban cities. As women are considered more vulnerable than men in urban landscapes, due to limited economic opportunities, poor workforce participation and less household standing, all mainly due to the patrilineal system. Therefore, policymakers in the government should work more on women-centric programs to challenge the patriarchal hold of men over assets and financial decision-making. Microfinance programs should be encouraged as they have been proven instrumental in improving social networks, mobility and self-employability, especially among women in traditional communities (Singh, 2015).

Notes

1. In the past couple of years, crimes against women have been steadily increasing. As per the National Crime Record Bureau (NCRB), India reported 3.59 lakh cases in 2017 up from 3.38 lakh cases in 2016 and 3.29 lakh cases in 2015 which includes rapes, domestic violence, acid attack, murder, dowry death, kidnapping and suicide abetment, etc. Available at https://ncrb.gov.in/ (accessed 01.06.2020).

2. In informal discussions with newly married couples living in Mumbai city of state Maharashtra reported that they need a separate house, television and other brown goods to setup the family and hence demand for more loan.
3. Dowry in India is seen as a transfer of money, gifts, goods, property, etc., from bride's side to groom's family at the time of marriage and sometimes after marriage too. Dowry culture is prominent in many north Indian states (Punjab, Haryana, Delhi, Uttar Pradesh, etc.).
4. Other background classes (OBC) is an official term coined by the Government of India to represent caste which is Educationally or Socially weaker sections (ESWS).

References

Ando, A. and Modigliani, F. (1963), "The "life cycle" hypothesis of saving: aggregate implications and tests", *The American Economic Review*, Vol. 53 No. 1, pp. 55-84, doi: 10.1126/science.151.3712.867-a.

Aterido, R., Beck, T. and Iacovone, L. (2011), *Gender and Finance in Sub-Saharan Africa: are Women Disadvantaged*?, The World Bank, Washington, DC.

Bag, S., Seth, S. and Gupta, A. (2016), "A comparative study of living conditions in slums of three metro cities in India", Centre for Development Economics (CDE) Working paper no. 253, ISSN 2454 – 1427, February, available at: www.cdedse.org/pdf/work253.pdf (accessed 20 August 2017).

Balasubramanian, S.A., Kuppusamy, T. and Natarajan, T. (2019), "Financial inclusion and land ownership status of women", *International Journal of Development Issues*, Vol. 18 No. 1, pp. 51-69, doi: 10.1108/IJDI-06-2018-0091.

Banerjee, A., Duflo, E., Glennerster, R. and Kinnan, C. (2015), "The miracle of microfinance? evidence from a randomized evaluation", *American Economic Journal: Applied Economics*, Vol. 7 No. 1, pp. 22-53, doi: 10.1257/app.20130533.

Banerjee, S. and Govil, D. (2018), "Bengali female migrants: Domestic workers in Mumbai", *ANTYAJAA: Indian Journal of Women and Social Change*, Vol. 3 No. 2, pp. 194-206.

Basargekar, P. (2010), "Measuring effectiveness of social capital in microfinance: a case study of urban microfinance programme in India", *International Journal of Social Inquiry*, Vol. 3 No. 2.

Batra, R. and Reio, T.G. (2016), "Gender inequality issues in India", *Advances in Developing Human Resources*, Vol. 18 No. 1, pp. 88-101, doi: 10.1177/1523422316630651.

Bennett, L. (1992), *Women, Poverty, and Productivity in India*, World Bank, Washington, DC.

Berger, M. (1989), "Giving women credit: the strengths and limitations of credit as a tool for alleviating poverty", *World Development*, Vol. 17 No. 7, pp. 1017-1032.

Bhalotra, S.R., Brulé, R. and Roy, S. (2017), "Women's inheritance rights reform and the preference for sons in India", IZA Discussion Papers, No. 11239, Institute of Labor Economics (IZA), Bonn.

Bloom, S.S., Wypij, D. and Gupta, M.D. (2001), "Dimensions of women's autonomy and the influence on maternal health care utilization in a North Indian city", *Demography*, Vol. 38 No. 1, pp. 67-78.

Bryman, A. and Bell, E. (2011), *Business Research Methods*, Oxford University Press.

Burgess, R. and Pande, R. (2005), "Do rural banks matter? Evidence from the Indian social banking experiment", *American Economic Review*, Vol. 95 No. 3, pp. 780-795.

Census of India (2011), "Number of households availing banking services and number of households having each of the specified assets, office of the registrar general census commissioner, India, Indian Census Bureau", available at: www.censusindia.gov.in/2011census/Hlo-series/HH12.html (accessed 9 October 2017).

Chadda, R. and Sinha, K. (2013), "Indian family systems, collectivistic society and psychotherapy", *Indian Journal of Psychiatry*, Vol. 55 No. 6, pp. S299-S309, doi: 10.4103/0019-5545.105555.

Chaudhuri, K. and Cherical, M.M. (2012), "Credit rationing in rural credit markets of India", *Applied Economics*, Vol. 44 No. 7, pp. 803-812.

Cochran, W.G. (1977), *Sampling Techniques*, 3rd ed. Wiley, New York, NY.

Crook, J. (2001), "The demand for household debt in the USA: evidence from the 1995 survey of consumer finance", *Applied Financial Economics*, Vol. 11 No. 1, pp. 83-91, doi: 10.1080/09603100150210291.

Datta, S. and Sahu, T.N. (2018), "Microcredit, employability and economic upliftment: a study on selected female borrowers in West Bengal", *The Management Accountant Journal*, Vol. 53 No. 11, pp. 71-77.

Davidson, R. and Mackinnon, J.G. (2004), *Econometric Theory and Methods*, Oxford University Press, New York, NY.

Deaton, A. (2005), "Franco modigliani and the life cycle theory of consumption", *BNL Quarterly Review*, Vol. 8 Nos No. 233/234, pp. 91-107, doi: 10.2139/ssrn.686475.

Delechat, M.C.C. Newiak, M.M. Xu, R. Yang, M.F. and Aslan, G. (2018), "What is driving women's financial inclusion across countries? International monetary fund.", doi: 10.5089/9781484344460.001

Demirguc-Kunt, A., Klapper, L., Singer, D., Ansar, S. and Hess, J. (2018), "*The Global Findex Database 2017 – Measuring Financial Inclusion and the Fintech Revolution*", The World Bank, Washington, DC, 10.1596/978-1-4648-1259-0.

Duflo, E. (2012), "Women's empowerment and economic development", *Journal of Economic Literature*, Vol. 50 No. 4, pp. 1051-1079, doi: 10.1177/0002716297554001009.

Dupas, P. and Robinson, J. (2013), "Savings constraints and microenterprise development: Evidence from a field experiment in Kenya", *American Economic Journal: Applied Economics*, Vol. 5 No. 1, pp. 163-192.

Feigenberg, B., Field, E., Pande, R., Rigol, N. and Sarkar, S. (2014), "Do group dynamics influence social Capital gains among microfinance clients? Evidence from a randomized experiment in urban India", *Journal of Policy Analysis and Management*, Vol. 33 No. 4, pp. 932-949.

Fidrmuc, J., Hake, M. and Stix, H. (2013), "Households' foreign currency borrowing in Central and Eastern Europe", *Journal of Banking and Finance*, Vol. 37 No. 6, pp. 1880-1897.

Fletschner, D. (2009), "Rural women's access to credit: market imperfections and intrahousehold dynamics", *World Development*, Vol. 37 No. 3, pp. 618-631.

Fletschner, D. and Carter, M.R. (2008), "Constructing and reconstructing gender: reference group effects and women's demand for entrepreneurial Capital", *The Journal of Socio-Economics*, Vol. 37 No. 2, pp. 672-693.

Fletschner, D. and Kenney, L. (2014), "Rural women's access to financial services: Credit, savings, and insurance", ESA Working Paper No. 11-07, The Food and Agriculture Organization of the United Nations. FAO, doi: 10.1007/978-94-017-8616-4_8.

Garcia, J.L., Digiannatale, S. and Carreon, V.G. (2015), "Why do formal credit, informal credit, and both types of credits coexist as consumer choices?", *Economics Bulletin*, Vol. 35 No. 1, pp. 89-96.

Garikipati, S. (2008a), "Redefining gender roles and reworking gender relations: Female agricultural labour in dry regions of Andhra Pradesh".

Garikipati, S. (2009), "Landless but not assetless: Female agricultural labour on the road to better status, evidence from India", *The Journal of Peasant Studies*, Vol. 36 No. 3, pp. 517-545, doi: 10.1080/03066150903142774.

Garikipati, S. (2017), "Microcredit and women's empowerment: through the lens of time-use data from rural India", *Development and Change*, Vol. 43 No. 3, pp. 719-750.

Ghosh, S. and Vinod, D. (2017), "What constrains financial inclusion for women? Evidence from Indian micro data", *World Development*, Vol. 92, pp. 60-81, doi: 10.1016/j.worlddev.2016.11.011.

Giné, X. (2011), "Access to xapital in rural Thailand: an estimated model of formal vs. informal credit", *Journal of Development Economics*, Vol. 96 No. 1, pp. 16-29.

Goedecke, J., Guérin, I., D'espallier, B. and Venkatasubramanian, G. (2018), "Why do financial inclusion policies fail in mobilizing savings from the poor? Lessons from rural South India", *Development Policy Review*, Vol. 36, pp. O201-O219.

Guérin, I. (2014), "Juggling with debt, social ties, and values: the everyday use of microcredit in rural South India", *Current Anthropology*, Vol. 55, pp. S40-S50, doi: 10.1086/675929.

Guérin, I., D'Espallier, B. and Venkatasubramanian, G. (2013), "Debt in rural South India: fragmentation, social regulation and discrimination", *Journal of Development Studies*, Vol. 49 No. 9, pp. 1155-1171, doi: 10.1080/00220388.2012.720365.

Harker, C., Sayyad, D. and Shebeitah, R. (2019), "The gender of debt and space: notes from Ramallah-Al Bireh, Palestine", *Geoforum*, Vol. 98, pp. 277-285.

Himabindu, B.L., Arora, R. and Prashanth, N.S. (2014), "Whose problem is it anyway? Crimes against women in India", *Global Health Action*, Vol. 7 No. 1, p. 23718.

Holvoet, N. (2005), "The impact of microfinance on decision-making agency: evidence from South India", *Development and Change*, Vol. 36 No. 1, pp. 75-102, doi: 10.1111/j.0012-155X.2005.00403.x.

Hosmer, D.W., Lemeshow, S. and Sturdivant, R.X. (1989), "The multiple logistic regression model", *Applied Logistic Regression*, Vol. 1, pp. 25-37.

Hung, A., Yoong, J. and Brown, E. (2012), "Empowering women through financial awareness and education", OECD Working Papers on Finance, Insurance and Private Pensions, No. 14, OECD Publishing, doi: 10.1787/5k9d5v6kh56g-en.

Iregui-Bohórquez, A.M., Melo-Becerra, L.A., Ramírez-Giraldo, M.T. and Tribín-Uribe, A.M. (2018), "Formal and informal household credit in Colombia", in Roa, M.J. and Mejía, D. (Eds), *Financial Decisions of Households and Financial Inclusion: Evidence for Latin America and the Caribbean*, Center for Latin American monetary studies, Mexico, pp. 125-153.

Islam, A., Nguyen, C. and Smyth, R. (2015), "Does microfinance change informal lending in village economies? Evidence from Bangladesh", *Journal of Banking and Finance*, Vol. 50, pp. 141-156, doi: 10.1016/j.jbankfin.2014.10.001.

Johanson, G.A. and Brooks, G.P. (2010), "Initial scale development: sample size for pilot studies", *Educational and Psychological Measurement*, Vol. 70 No. 3, pp. 394-400.

Kabeer, N. (2001), "Conflicts over credit: Re-evaluating the empowerment potential of loans to women in rural Bangladesh", *World Development*, Vol. 29 No. 1, pp. 63-84, doi: 10.1016/S0305-750X(00)00081-4.

Kaboski, J.P. and Townsend, R.M. (2012), "The impact of credit on village economies", *American Economic Journal: Applied Economics*, Vol. 4 No. 2, pp. 98-133.

Kar, S. (2018), "Securitizing women: gender, precaution, and risk in Indian finance", *Signs: Journal of Women in Culture and Society*, Vol. 43 No. 2, pp. 301-325.

Karim, L. (2011), *Microfinance and Its Discontents, Women in Debt in Bangladesh, Minneapolis and London*, University of MN Press.

Kaushik, N., Sharma, A. and Kaushik, V.K. (2014), "Equality in the workplace: a study of gender issues in Indian organisations", *Journal of Management Development*, Vol. 33 No. 2, pp. 90-106.

Koride, P. and Gurtoo, A. (2019), "A comparison of borrowing and default behaviour between men and women", *Studies in Microeconomics*, Vol. 7 No. 2, pp. 189-209.

Lee, S. and Persson, P. (2016), "Financing from family and friends", *Review of Financial Studies*, Vol. 29 No. 9, pp. 2341-2386, doi: 10.1093/rfs/hhw031.

Li, S., Liu, Y. and Deininger, K. (2013), "How important are endogenous peer effects in group lending? Estimating a static game of incomplete information", *Journal of Applied Econometrics*, Vol. 28 No. 5, pp. 864-882, doi: 10.1002/jae.2276.

Malapit, H.J.L. (2012), "Are women more likely to be credit constrained? Evidence from low-income urban households in the Philippines", *Feminist Economics*, Vol. 18 No. 3, pp. 81-108.

Mallick, D. (2012), "Microfinance and moneylender interest rate: evidence from Bangladesh", *World Development*, Vol. 40 No. 6, pp. 1181-1189, doi: 10.1016/j.worlddev.2011.12.011.

Mohiuddin, Y. and Al-Jalaly, S.Z. (1993), "Credit worthiness of poor women: a comparison of some minimalist credit programmes in Asia: a preliminary analysis", *The Pakistan Development Review*, Vol. 32 No. 4II, pp. 1199-1209.

Mpuga, P. (2010), "Constraints in access to and demand for rural credit: Evidence from Uganda", *African Development Review*, Vol. 22 No. 1, pp. 115-148, doi: 10.1111/j.1467-8268.2009.00230.x.

Nguyen, C.H. (2007), "Access to credit and borrowing behaviour of rural households in a transition economy", *International Conference on Rural Finance Research: Moving Results into Policies and Practice*, FAO, Rome, 19-21.

Nieder, C., Muck, C. and Kärtner, J. (2019), "Sexual violence against women in India: Daily life and coping strategies of young women in Delhi", *Violence against Women*, Vol. 25 No. 14, pp. 1717-1738.

Noponen, H. (1992), "Loans to the working poor: a longitudinal study of credit, gender and the household economy", *International Journal of Urban and Regional Research*, Vol. 16 No. 2, pp. 234-251.

Pal, S. (2002), "Household sectoral choice and effective demand for rural credit in India", *Applied Economics*, Vol. 34 No. 14, pp. 1743-1755, doi: 10.1080/00036840210121228.

Parappurathu, S., Ramachandran, C., Baiju, K.K. and Xavier, A.K. (2019), "Formal versus informal: Insights into the credit transactions of small-scale fishers along the South West Coast of India", *Marine Policy*, Vol. 103, pp. 101-112.

Pastrapa, E. and Apostolopoulos, C. (2015), "Estimating determinants of borrowing: Evidence from Greece", *Journal of Family and Economic Issues*, Vol. 36 No. 2, pp. 210-223, doi: 10.1007/s10834-014-9393-4.

Pitt, M.M., Khandker, S.R. and Cartwright, J. (2006), "Empowering women with micro finance: evidence from Bangladesh", *Economic Development and Cultural Change*, Vol. 54 No. 4, pp. 791-831, doi: 10.1086/503580.

Postelnicu, L., Hermes, N. and Szafarz, A. (2014), "Defining social collateral in microfinance group lending", in Mersland, R., Strom R.O. (Eds), *Microfinance Institutions, Palgrave Studies in Impact Finance*, Palgrave Macmillan, London, doi: 10.1057/9781137399663.

Rahji, M.A.Y. and Fakayode, S.B. (2009), "A multinomial logit analysis of agricultural credit rationing by commercial banks in Nigeria", *International Research Journal of Finance and Economics*, Vol. 24 No. 91, pp. 97-103.

Rao, N. (2017), *Good Women Do Not Inherit Land': Politics of Land and Gender in India*, Routledge.

Rawat, P.S. (2014), "Patriarchal beliefs, women's empowerment, and general well-being", *Vikalpa: The Journal for Decision Makers*, Vol. 39 No. 2, pp. 43-56.

Razvi, M. and Roth, G.L. (2004), *Socio-Economic Development and Gender Inequality in India*, Online Submission, Northern IL University, available at: https://files.eric.ed.gov/fulltext/ED492144.pdf (accessed 2 January 2020).

Robitaille, M.C. (2020), "Conspicuous daughters: Exogamy, marriage expenditures, and son preference in India", *The Journal of Development Studies*, Vol. 56 No. 3, pp. 630-647.

Ruthven, O. (2002), "Money mosaics: Financial choice and strategy in a West Delhi squatter settlement", *Journal of International Development*, Vol. 14 No. 2, pp. 249-271.

Sahay, M.R. and Cihak, M.M. (2018), "Women in finance: a case for closing gaps, international monetary fund.", *Staff Discussion Notes*, Vol. 18 No. 05.

Sankaran, K. and Madhav, R. (2011), *Gender Equality and Social Dialogue in India*, International Labour Office, ILO, Geneva.

Sanyal, P. (2014), *Credit to Capabilities: A Sociological Study of Microcredit Groups in India*, Cambridge University Press, Cambridge.

Sim, J. and Lewis, M. (2012), "The size of a pilot study for a clinical trial should be calculated in relation to considerations of precision and efficiency", *Journal of Clinical Epidemiology*, Vol. 65 No. 3, pp. 301-308.

Singh, S. (2015), "The effects of microfinance programs on women members in traditional societies", *Gender, Place & Culture*, Vol. 22 No. 2, pp. 222-238.

Sinha, M., Mahapatra, S.S., Dutta, A. and Sengupta, P.P. (2019), "Microfinance and women empowerment: an empirical analysis", *Handbook of Research on Microfinancial Impacts on Women Empowerment, Poverty, and Inequality*, IGI Global, pp. 52-64.

Straub, S. (2005), "Informal sector: the credit market channel", *Journal of Development Economics*, Vol. 78 No. 2, pp. 299-321, doi: 10.1016/j.jdeveco.2004.09.005.

Swain, R.B. (2002), "Credit rationing in rural India", *Journal of Economic Development*, Vol. 27 No. 2, pp. 1-20.

Swain, R.B. and Wallentin, F.Y. (2012), "Factors empowering women in Indian self-help group programs", *International Review of Applied Economics*, Vol. 26 No. 4, pp. 425-444.

Swamy, V. (2014), "Financial inclusion, gender dimension, and economic impact on poor households", *World Development*, Vol. 56, pp. 1-15, doi: 10.1016/j.worlddev.2013.10.019.

Vicente, G., Goicoa, T., Fernandez-Rasines, P. and Ugarte, M.D. (2020), "Crime against women in India: unveiling spatial patterns and temporal trends of dowry deaths in the districts of Uttar Pradesh", *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, Vol. 183 No. 2, pp. 655-679.

Wang, F. (2019), "An empirical equilibrium model of formal and informal credit markets in developing countries", SSRN, January, available at: https://ssrn.com/abstract=3316939 (accessed 2 May 2019).

Wu, J. (2018), "The coexistence and interaction of formal and informal lending in China – discussion of the Wenzhou case", *The Chinese Economy*, Vol. 51 No. 1, pp. 97-114.

Yuan, Y. and Gao, P. (2012), "Farmer's financial choices and informal credit markets in China", *China Agricultural Economic Review*, Vol. 4 No. 2, pp. 216-232, doi: 10.1108/17561371211224791.

Further reading

Chen, K.C. and Chivakul, M. (2008), "What drives household borrowing and credit constraints? Evidence from Bosnia and Herzegovina", IMF Working Paper No. 8-202, International Monetary Fund.

Garikipati, S. (2008b), "The impact of lending to women on household vulnerability and women's empowerment: evidence from India", *World Development*, Vol. 36 No. 12, pp. 2620-2642, doi: 10.1016/j.worlddev.2007.11.008.

Prina, S. (2015), "Banking the poor via savings accounts: Evidence from a field experiment", *Journal of Development Economics*, Vol. 115, pp. 16-31.

Swain, R.B. (2007), "The demand and supply of credit for households", *Applied Economics*, Vol. 39 No. 21, pp. 2681-2992.

Corresponding author

Shruti Malik can be contacted at: shrutimalik.2050@gmail.com

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**

Environmental Management
https://doi.org/10.1007/s00267-020-01361-x

Hydrogeochemical Assessment of Groundwater for Drinking and Agricultural Use: A Case Study of Rural Areas of Alwar, Rajasthan

Ankur Aggarwal[1] · Jigyasa Soni[1] · Khyati Sharma[1] · Mohnish Sapra[1] · Chitrakshi[1] · Oznur Karaca[2] · A. K. Haritash[1]

Received: 6 April 2020 / Accepted: 5 September 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract
Groundwater contributes substantially to the development of arid and semi-arid regions around the globe. The present study integrates groundwater quality and its suitability for drinking and irrigation around Alwar city of Rajasthan state, where agriculture is the major land use. The application for drinking was assessed by comparing the observed value with prescribed standards of WHO. Groundwater was found suitable for drinking at most of the locations. The suitability of groundwater for irrigation was determined by calculating ion-based ratios and comparing them against the suggested ratios and indices for agricultural quality. Suitability for irrigation was assessed against electrical conductivity (EC), percentage sodium (%Na), residual Na_2CO_3 (RSC), per cent soluble sodium (SSP), sodium adsorption ratio (SAR), Mg hazard and permeability index (PI) etc., and the quality was compromised for EC, %Na and Mg Hz. Since the soil was sandy, the groundwater was found suitable for irrigation over long-term use, with the only problem of magnesium hazard. Based on the different ratios of anions and cations, silicate weathering was observed to be regulating groundwater chemistry, and the groundwater belonged to mixed CaMgCl and $CaHCO_3^-$ type based on Piper's classification and relative abundance of ions. Further, meteoric genesis classification showed that the groundwater in the study region had direct base exchange and shallow meteoric water percolation. Presence of kaolinite and quartz minerals in soil confirmed that silicate weathering is the major process controlling groundwater chemistry.

Keywords Siliserh Lake · Groundwater · Irrigation · Silicate weathering · FTIR

Introduction

Groundwater is the source of daily water supply to about 3.0 billion people worldwide. Of the total groundwater extraction globally, about 40% is used in agriculture (Haritash et al. 2017; Kumar et al. 2017). Agriculture is a major economic sector in India contributing to 46% of the gross national product (Kumarasamy et al. 2011). Since the advent of intensive agriculture, the pressure on groundwater has increased to meet the demand for irrigation. As a result, the groundwater table in many regions in India is receding, and issues of water quality have emerged particularly in water-scarce regions. Groundwater plays an important role in agricultural and industrial production around the globe, and about 90% of the rural population in India is dependent on it (Haritash et al. 2008). Groundwater quality has been deteriorating over time due to the rapid growth of population, urbanisation, industrialisation and land-use change. Other factors having an indirect impact on the quantity and quality of groundwater include climate change, poor management and limited regulations for groundwater. Arid and semi-arid regions which experience low rainfall and high evaporation rates are significantly affected, resulting in depletion of the groundwater table. Over extraction of groundwater is another reason for the receding groundwater table in many regions of the world. Since India has fragmented agricultural land and almost every farmer has a borewell to extract water for irrigation, excessive extraction of groundwater is being reported. Based on physical testing, groundwater often appears clean and clear since it gets filtered through the soil-bed, but natural and anthropogenic activities may deteriorate the groundwater quality. Rock–water interaction, weathering, stormwater runoff, agricultural runoff, use of fertilisers and pesticides for crop production, improper waste management and industrial

✉ A. K. Haritash
akharitash@dce.ac.in

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

[2] Department of Geological Engineering, Çanakkale Onsekiz Mart Üniversity, Canakkale, Turkey

Published online: 16 September 2020

processes are among the natural and anthropogenic process/activities, which cause groundwater contamination (Chitrakshi and Haritash 2018; Karaca et al. 2017). Poor well construction and inadequate groundwater well covers are other sources of direct contamination.

The present study is the first comprehensive effort to investigate groundwater quality and its suitability for use in the villages around Siliserh Lake, Rajasthan. Rajasthan is the largest state of India with an area of 342,239 km^2 and the Alwar district of Rajasthan lies in the semi-arid zone. Annual average rainfall varies from 595 mm to as low as 100 mm in the north-west region of Rajasthan and the climate is sub-humid. The average depth of the water table is around 40 m with an annual fluctuation of more than 4 m (Central Ground Water Board 2016). The groundwater is recharged by rainfall during the monsoon but is significantly depleted during the winter and summer. Although precipitation helps in recharging the groundwater, it also contaminates the groundwater by surface runoff, domestic runoff and agricultural runoff resulting in increased concentrations of nitrates, phosphates, Cl^-, microbes or synthetic chemicals (Food and Agriculture Organisation 2017). The agricultural land around the Siliserh Lake is often submerged during the monsoon, and during summers the water level of the lake falls leaving behind a high concentration of salts, which decreases soil productivity. Siliserh Lake also attracts migratory birds including cranes, and bird droppings along with agricultural runoff may further elevate the concentration of phosphates in groundwater after percolation.

Considering these problems, the study was conducted in five villages, around the Siliserh lake environs, namely Paitpur, Kishanpura, Bhagatpura, Doba and Ringaspura. Tube wells, open wells and handpumps in the study area are being used for groundwater extraction for drinking, domestic use and irrigation. The objectives of the present study were to generate the background data on quality of groundwater around the Siliserh Lake and to evaluate the suitability of groundwater for drinking and irrigation. Since the groundwater chemistry is controlled by rock–water interaction and anthropogenic inputs, the geochemical analysis of ions present in groundwater was performed to identify the major sources responsible for regulating groundwater chemistry.

Materials and Methods

Alwar is located at 27°02′38″ North latitude and 74°21′79″ East longitude in the northeastern part of Rajasthan state. The district has 8720 km^2 area with a population of 3.6 million as per population census of year 2011. Siliserh Lake is the major source for drinking water supply, agricultural use and livestock management in villages namely, Paitpur, Kishanpura, Bagatpura, Doba and Ringaspura located around the lake. Residents of these villages are dependent on the groundwater for drinking and irrigation. Since land use in this region is mostly agricultural, it was important to analyse the quality of groundwater for drinking as well as irrigation. Groundwater samples were collected from twenty-three locations covering the villages located on the periphery of the Siliserh Lake (Fig. 1). The groundwater samples were collected from tube wells, handpumps and open wells during April 2018. The pre-rinsed polypropylene plastic bottles were used for the collection of samples and were stored at the low temperature (4 °C) till chemical analysis. The coordinates of each location were marked and details of site-specific activities were noted. The physicochemical analysis of the sampled groundwater was carried out within the 24–48 h after collection. The pH, electrical conductivity (EC), total dissolved salts (TDS) and temperature were measured onsite using 'Orion make Star A320 model multi-parameter meter' and rest of the parameters were analysed in triplicates using AR grade reagents and type 1 ultra-pure water as per the standard methods (APHA 2005). Total alkalinity, total hardness (TH), chloride (Cl^-), carbonate (CO_3^{2-}), bicarbonate (HCO_3^-) and silicate were determined by volumetric analysis; and PO_4^{3-}, SO_4^{2-}, NO_3^- and Fe were analysed using UV–Vis spectrophotometer (LabIndia make UV3092 model). Concentrations of Na^+, K^+, Ca^{2+} and Mg^{2+} were determined using Systronics make 128μC model Flame photometer. The suitability of groundwater for drinking was analysed by comparing the average value with the prescribed standards of Bureau of Indian Standards (2012), and by calculating water quality index (WQI) for its classification into different categories. WQI was calculated using the following equation:

$$WQI = \sum WiQi / \sum Wi,$$

where, Q is the quality rating for ith parameter [$Q_i = 100 \times (V_i - V_o)/(S_i - V_o)$]; V_i is the observed value of the parameter; V_O is the ideal value of the parameter and S_i is the standard value of the parameter. For determining agricultural suitability, various indices and ratios were calculated as discussed elsewhere (Chitrakshi and Haritash 2018). The parameters used for assessing the suitability for irrigation were EC, percentage sodium (%Na), residual Na_2CO_3 (RSC), per cent soluble sodium, sodium adsorption ratio (SAR), Mg hazard and permeability index (PI). For the geochemical analysis, various ratios (Ca/Na; Mg/Na; HCO_3^-/Na) (Gaillardet et al. 1999) were calculated to ascertain the type of weathering controlling the groundwater chemistry. To determine the type of water, Piper tri-linear classification (Piper 1953) was used by plotting the diagram

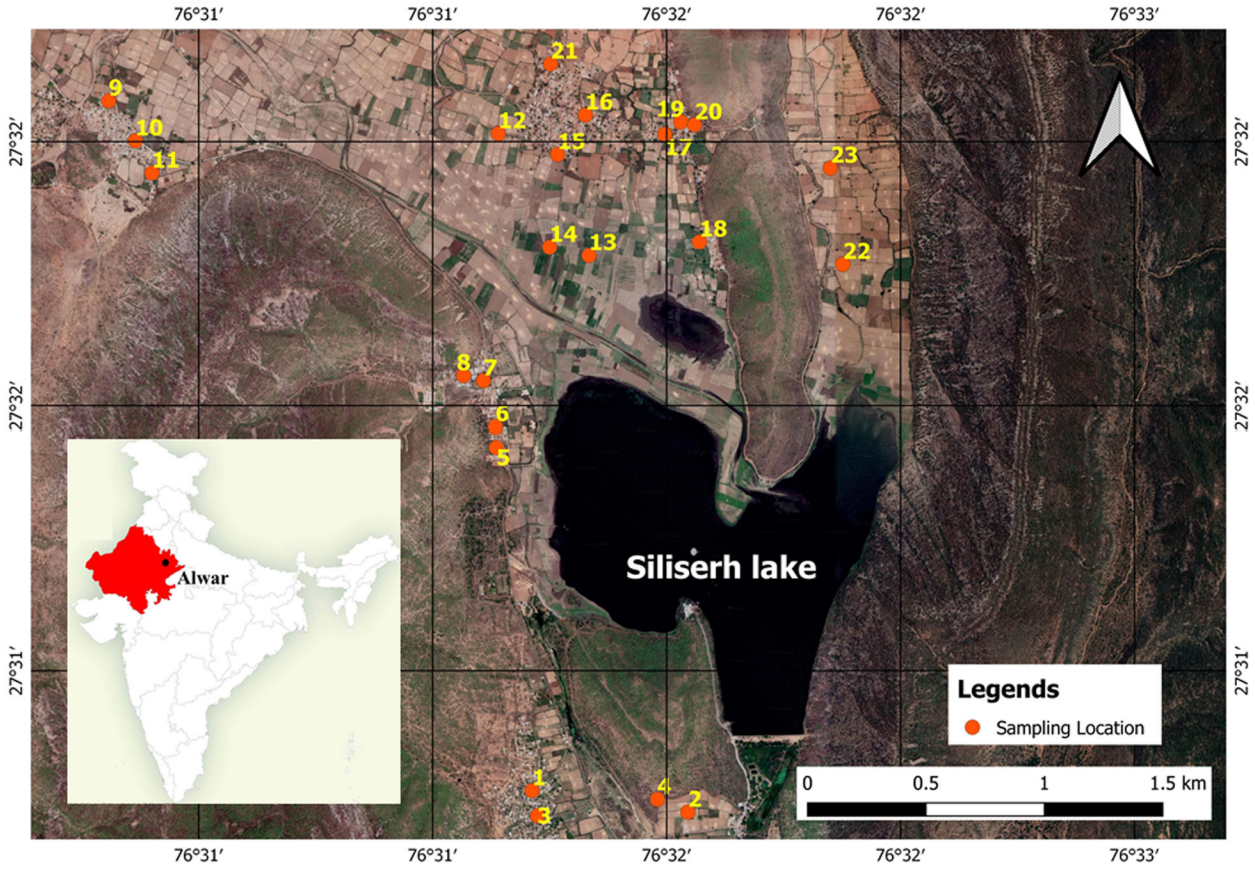


Fig. 1 Location of sampling points around Siliserh Lake Region, Alwar, Rajasthan (picture in the left bottom corner shows the location of Alwar in India)

over 'GW-Chart software of USGS'. The base-exchange (BE) index for the groundwater was also calculated using the equation $(Na^{+}–Cl^{-})/SO_4^{2-}$. Further, FTIR analysis of soil collected from the agricultural fields was performed to determine the chief minerals present in it.

Results and Discussion

A notable concentration of all the major cations and anions was observed in the samples upon its physicochemical analysis. The results so obtained are presented in Table 1. Average pH of the analysed samples was near neutral with a value of 6.8. The slightly alkaline behaviour of water shows the dominance of alkali ions (Ca^{2+} and Mg^{2+}) over the acidic ions (SO_4^{2-} and Cl^{-}). The EC varied from 586 to 4300 µS/cm with an average value of 1209 µS/cm. High values of EC in the groundwater may be attributed to high salinity and high TDS in water. Relatively higher values of EC were found in the samples taken from the open wells of Doba village (sample number 13, 14, 15). Similar to EC, high values of TDS were observed at these locations due to higher amounts of dissolved Ca, Na, Mg, HCO_3, Cl^{-} and SO_4^{2-} in groundwater.

Suitability of Sampled Water for Drinking in Rural Areas of Alwar, Rajasthan

The results obtained during analysis were compared with standards specified by Bureau of Indian Standards (2012) and World Health Organization (2006) for determining the suitability for drinking (Table 2). In this study, TDS varied from 276 to 2150 mg/l with an average value of 586 mg/l. Upon comparison with the prescribed standards of drinking water (Bureau of Indian Standards 2012), about 35% of the analysed samples were exceeding the standard concentration of 500 mg/l. Based on the US Geological Survey (2000) criterion, about 13% of the sampled water collected from open wells of Doba was categorised as saline. The average value of TH (TH as $CaCO_3^{2-}$) was found to be 390 mg/l, varying from 80 to 964 mg/l, with all samples exceeding the prescribed limit of 200 mg/l excluding one sample collected from the Paitpur village. Based on the classification of Sawyer et al. (2003), 4% of the samples were moderately hard; 35% of the samples were hard and 61% samples very hard, thus not fit for drinking and washing clothes. Based on the observed values of cations, Na was most abundantly present and

Table 1 Physicochemical characteristic of groundwater of rural areas in Alwar during April 2018

S. No.	Village	pH	EC (μS/cm)	TDS (mg/l)	TH (mg/l)	TA (mg/l)	Na^+ (mg/l)	K^+ (mg/l)	Ca^{2+} (mg/l)	Mg^{2+} (mg/l)	Li^+ (mg/l)	Iron (mg/l)	CO_3^{2-} (mg/l)	HCO_3^- (mg/l)	Cl^- (mg/l)	SO_4^{2-} (mg/l)	NO_3^- (mg/l)	$(PO_4)^{3-}$ (mg/l)
1	Paitpur	7.9	778	389	80	400	138	1.3	45	1.0	0.10	0.13	14	200	20	6.8	69	0.0
2	Paitpur	7.0	718	358	272	272	29	0.9	21	26.2	0.00	0.10	48	20	68	7.5	20	0.1
3	Paitpur	7.1	595	298	264	216	24	1.2	24	24.3	0.10	0.07	10	102	52	5.4	18	0.0
4	Paitpur	7.1	822	409	304	288	47	1.1	29	37.9	0.10	0.04	19	117	48	14.5	24	0.0
5	Kishanpura	6.8	1355	678	440	264	82	2.8	52	57.3	0.15	0.09	24	88	140	94.1	42	0.0
6	Kishanpura	7.0	1102	547	300	200	19	67.4	22	27.2	0.15	0.02	19	63	92	54.9	49	0.9
7	Kishanpura	6.7	643	322	244	136	21	12.8	22	18.5	0.10	0.11	0	83	52	16.3	22	0.0
8	Kishanpura	6.3	803	399	252	176	16	37.1	20	23.3	0.10	0.05	10	78	60	54.4	22	0.0
9	Bakhtpura	6.9	676	337	292	224	12	3.9	22	24.3	0.05	0.06	0	137	40	8.2	22	0.1
10	Bakhtpura	7.1	1604	803	624	224	42	9.2	77	71.8	0.15	0.11	5	122	188	139.3	59	0.0
11	Bakhtpura	6.8	768	384	224	216	14	3.6	25	0.0	0.05	0.09	0	132	36	11.7	18	0.0
12	Doba	6.8	1225	276	456	264	62	1.8	48	53.4	0.10	0.03	0	161	140	101.3	27	0.2
13	Doba	6.6	2190	1097	768	400	167	6.7	92	101.9	0.20	0.02	0	244	276	168.9	50	0.2
14	Doba	6.8	2400	1203	740	368	243	1.7	101	100.9	0.15	0.03	0	224	356	145.1	55	0.0
15	Doba	6.8	4300	2150	964	392	377	1.5	141	142.7	0.25	0.26	0	239	852	203.8	78	0.0
16	Doba	6.9	1372	683	540	336	73	3.3	48	68.9	0.10	0.02	0	205	132	59.8	33	0.2
17	Doba	6.9	762	381	232	200	32	1.5	34	22.3	0.05	0.06	0	122	52	23.6	24	0.1
18	Ringaspura	6.7	834	339	320	256	42	8.7	35	41.7	0.05	0.15	0	156	60	15.8	16	0.2
19	Ringaspura	7.3	1001	499	368	288	55	8.4	41	60.2	0.10	0.08	0	176	76	44.2	20	0.2
20	Ringaspura	6.7	1632	814	320	208	54	3.5	66	0.0	0.25	0.41	0	127	244	87.1	40	0.0
21	Ringaspura	6.8	933	467	320	280	68	4.4	49	25.2	0.10	0.44	0	171	68	25.5	26	0.1
22	Ringaspura	6.7	708	353	276	216	26	1.4	28	10.7	0.05	0.09	0	132	48	15	19	0.1
23	Ringaspura	6.6	586	293	364	208	23	1.7	32	62.1	0.05	0.13	0	127	32	7.2	12	0.1
Mean		6.8	1209	586	390	262	72	8.0	47	42	0.11	0.11	6.5	140	136	57	33	0.10
SD (±)		0.31	835	425.33	208	73.72	86.51	15.01	30.53	38.44	0.06	0.11	11.81	57.12	178.57	59.01	18.44	0.19

Table 2 Per cent suitability of groundwater for drinking against the prescribed standards

S. No.	Parameters	Range (min–max)	Mean ± SD	Standard for drinking		Suitability (%)
				BIS	WHO	
1	pH	6.3–7.9	6.80 ± 0.31	6.5–8.5	8.5	100
2	EC	586–4300	1209 ± 835	NM*	NM	
3	TDS (mg/l)	276–2150	586 ± 425	500	500	65
5	TH (mg/l)	80–1964	390 ± 208	200	500	4
6	TA (mg/l)	136–400	262 ± 73.72	200	250	9
7	CO_3^{2-} (mg/l)	0–19.2	6.47 ± 11.81	NM	NM	–
8	HCO_3^- (mg/l)	20–239	140 ± 57.12	NM	NM	–
9	Cl^- (mg/l)	20–852	136 ± 178.57	250	200	87
10	$S0_4^{2-}$(mg/l)	5.4–203.8	57 ± 59.01	200	250	96
11	NO_3^- (mg/l)	12–78	33 ± 18.44	45	11	74
12	Na^+ (mg/l)	12–377	72 ± 86.51	NM	NM	–
13	K^+ (mg/l)	0.85–67.4	8 ± 15.01	NM	NM	–
14	Ca^{2+} (mg/l)	19.90–140.65	47 ± 30.53	75	75	82
15	Mg^{2+} (mg/l)	0–142.72	42 ± 38.44	30	50	52
16	Li^{2+} (mg/l)	0–0.25	0.11 ± 0.06	NM	NM	–
17	Fe (mg/l)	0.021–0.44	0.11 ± 0.11	0.3	0.3	91
18	WQI	7–138	36 ± 35	0–25: Excellent 26–50: Good 51–75: Poor 76–100: Very poor >100: Unsuitable		52 35 00 04 09

*NM-not mentioned

followed by Ca, Mg, K and Li. The average value of Na was 72 mg/l, ranging from 12 to 377 mg/l. A high concentration of sodium at some locations in Paitpur and Doba villages may be attributed to high salinity in the groundwater at that location. The average value of Ca in the groundwater was 46, varying between 21 and 140 mg/l, with 82% of the collected samples below the permissible limit of 75 mg/l. Similarly, the average value of Mg was found to be 42 mg/l and upon comparison, 48% of the samples were exceeding the prescribed standard of 30 mg/l. The high levels were mainly observed in the open wells of Doba village. For Fe, only 9% of the samples were observed to be slightly above the prescribed limit of 0.3 mg/l; and others within the safe limit. Among the anions, Cl^- was the most abundant ion followed by SO_4^{2-}, NO_3^- and PO_4^{3-}. The concentration of Cl^- exceeded the prescribed standard in about 17% of the samples, while NO_3^- was in the range of 12–78 mg/l with a mean value of 33 mg/l. The relatively higher concentration of nitrate observed at Paitpur village (samples 1 and 6), Kishanpura (sample 10) and Doba village (samples 13, 14, 15) with the maximum value of 78 mg/l, may be attributed to the agricultural activities at these locations. SO_4^{2-} was found significantly below the prescribed BIS standard of 200 mg/l except for the sample 15 collected from the Doba village, which had a slightly higher value. Similarly, PO_4^{3-} also represented lower concentration in the groundwater of the region. The samples collected from the open wells of Doba village (13, 14 and 15), which were surrounded by the agricultural fields represented higher concentrations for most of the parameters. These wells get submerged annually during the rainy season, as reported by the villagers. A higher concentration of salts and other pollutants in this area may also be attributed to a higher rate of evaporation during summer since Alwar lies in a semi-arid region. Simultaneously, agricultural and domestic runoff percolates to the groundwater playing a role in contaminating the groundwater. Based on WQI, 86% of the samples representing good quality can be used for drinking after disinfection. Meanwhile, sampled water from the Doba and Ringaspura tube wells (samples 17 and 20) was unsuitable for drinking due to high WQI rating.

Suitability of Sampled Water for Irrigation in Rural Areas of Alwar, Rajasthan

Since major land-use in the study area is agricultural, it was important to determine the suitability of groundwater for irrigation. Various ion-based ratios/indices were estimated (Table 3) and compared to the prescribed standards

Table 3 Suitability of groundwater for irrigation based on different ratios/indices

S. No	SAR	RSC	SSP	%Na	Mg Hz	K.R.	P.I.
1	5.55	1.43	71.98	254.82	3.47	2.57	93.73
2	1.00	−1.32	28.28	39.83	67.38	0.39	40.79
3	0.82	−1.22	24.46	33.03	62.91	0.32	54.92
4	1.35	−2.06	30.81	44.85	68.26	0.45	51.55
5	1.85	−5.15	32.54	48.73	64.59	0.48	43.49
6	0.65	−1.68	19.93	50.39	67.36	0.25	44.21
7	0.78	−1.29	25.22	41.02	58.01	0.34	58.13
8	0.58	−1.34	19.34	42.56	66.12	0.24	50.42
9	0.40	−0.88	13.81	18.60	64.83	0.16	55.15
10	0.82	−7.66	15.62	20.41	60.97	0.19	27.77
11	0.75	0.92	32.37	51.43	0.00	0.48	112.53
12	1.46	−4.19	28.28	39.84	65.15	0.39	45.34
13	2.83	−9.10	35.61	55.89	64.87	0.55	45.45
14	4.08	−9.77	44.01	78.67	62.54	0.79	51.99
15	5.32	−15.01	46.39	86.57	62.84	0.87	52.00
16	1.57	−4.77	28.00	39.50	70.62	0.39	44.22
17	1.04	−1.57	28.02	39.59	52.08	0.39	56.51
18	1.12	−2.67	25.76	37.35	66.56	0.35	48.48
19	1.27	−4.16	25.26	35.75	71.22	0.34	43.27
20	1.83	−1.21	41.69	72.24	0.00	0.71	67.25
21	1.95	−1.77	39.19	65.30	45.99	0.64	61.44
22	1.04	−0.14	32.59	49.13	38.65	0.48	75.62
23	0.54	−4.67	12.76	15.15	76.71	0.15	31.40
Mean	1.68	−3.45	30.52	54.81	54.83	0.52	54.59

(Table 4) (Eaton 1950; Richards 1954; Wilcox 1955; Doneen 1964; Kelly 1963; Paliwal 1972) to determine its suitability for irrigation. EC is the direct inference of TDS in water, which affects the soil and the crop productivity. In the present study, 26% of the samples were classified as good, 62% were classified as moderate, 8% samples were poor; while the remaining 4% were not fit for irrigation. Increase of electrical conductivity can be correlated to reduced water intake by the plants resulting in lower crop production. Considering Na, 79% of the samples were classified as suitable (very good to moderate quality); and 21% were classified as unsuitable (poor quality of water). SAR is another Na-based parameter to determine groundwater suitability for irrigation because sodium can replace adsorbed calcium and magnesium thereby causing damage to the soil morphology, increased compaction and reduced permeability (Arveti et al. 2011). Based on SAR, all the samples were classified as fit for irrigation over long-term use. The average value of RSC was negative for all the samples, which indicate that Ca^{2+} and Mg^{2+} ions are more than Na^{+} ions, as a result of which there will not be any binding of Na and HCO_3^{-} ions resulting in residual sodium carbonate. However, for Mg hazard, only 17% of the samples were suitable for irrigation and 83% were observed unsuitable. For KR, all the samples were classified as suitable and can be used for irrigation over a long period. Similarly, all the samples were categorised in class I type, based on PI, which indicates that the groundwater quality is fit for irrigation.

Geochemical Characteristics of Sampled Water in Rural Areas of Alwar, Rajasthan

The major processes affecting the water chemistry are well described by Gibb's boomerang plot (Gibbs 1970) (Fig. 2). In the present study, rock–water interaction was observed to be the dominant process controlling the relative concentration of cations and anions in groundwater. Based on Ca/Na, Mg/Na and HCO_3^{-}/Na ratios, the values were found close to the silicate weathering and far below the CO_3^{2-} weathering. FTIR analysis, too, revealed the presence of kaolinite and quartz minerals in the soil (Fig. 3). Since kaolinite and quartz are the end member of the silicate weathering, it was concluded that the dominant process governing groundwater quality is the rock–water interface. The average value of $Na^{+}K/TZ^{+}$ close to 0.33 also confirms the dominance of silicate weathering over CO_3^{2-} weathering. Based on Piper-tri-linear plot (Fig. 4) of the major cations and anions in the groundwater, most of the groundwater samples fit mixed CaMgCl and $CaHCO_3^{-}$ type. Among the cations, Ca and Mg ion concentration were more than Na and K ions; while Cl^{-} dominates over other anions. According to the classification of meteoric genesis (MG) as discussed by Abdel Moneim (1988), a shallow meteoric water percolation type was observed with the average value of 1.12 for MG. The value of BE less than one indicates the source of Na-SO_4^{2-} type and the value of BE more than one indicates the water to be Na-HCO_3^{-} type. Based on the values, only 13% of the water samples were observed to be Na-HCO_3^{-} type and 87% of the samples were of Na-SO_4^{2-} type.

Conclusion

Based on the physicochemical characterisation, it is suggested that the groundwater around the Siliserh Lake is suitable for human consumption after conventional treatment. The groundwater in villages (Paitpur, Kishanpura and Doba) located north of Siliserh Lake had a relatively higher concentration of dissolved ions. Since the water flows into Siliserh from north-west region, which is dominantly agricultural land use, and dissolution of ions from interacting agricultural soil might be higher. It stresses the need for limiting agricultural activity and application of agrochemicals particularly in the northern

Table 4 Per cent suitability of water for irrigation against various indices

S. No.	Parameter	Mean value (Range: min–max)	Category	Inference	Suitable samples (%)
1	EC	1209 (586–4300)	250[a]	Very good	0
			250–750	Good	26
			750–2000	Marginal	62
			2000–3000	Very poor	8
			3000	Harmful	4
2	Na (%)	55 (15–255)	20[a]	Very good	9
			20–40	Good	35
			40–60	Marginal	35
			60–80	Poor	12
			80	Harmful	9
3	RSC	−3.45 (−15–1.43)	<0[b]	Very good	91
			0–2.5	Good	9
			2.5–5.0	Marginal	0
			5.0–7.5	Poor	0
			>7.5	Harmful	0
4	SAR	1.68 (0.40–5.55)	0–10[c]	Very good	100
			10–18	Good	0
			18–26	Marginal–poor	0
			>26	Harmful	0
5	Mg hazard	54.83 (0–77)	<50%[d]	Good	17
			>50%	Harmful	83
7	KR	0.52 (0.15–2.57)	<1[e]	Good–very good	96
			1–2	Marginal	4
			>2	Poor–harmful	0
8	PI	55 (28–113)	Class I	Safe	6
			Class II	Moderate	63
			Class III	Doubtful	31

[a]Wilcox 1955; [b]Bishnoi et al. 1984; [c]Richards 1954; [d]Paliwal 1972; [e]Kelly 1963

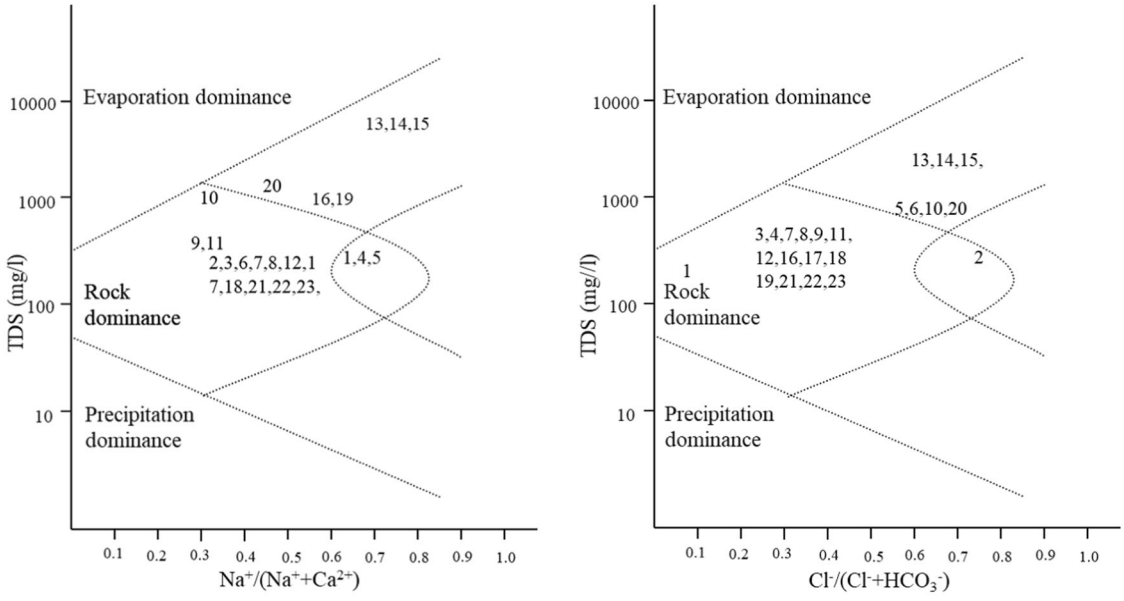


Fig. 2 Gibbs boomerang plot for mechanism governing water chemistry around Siliserh Lake

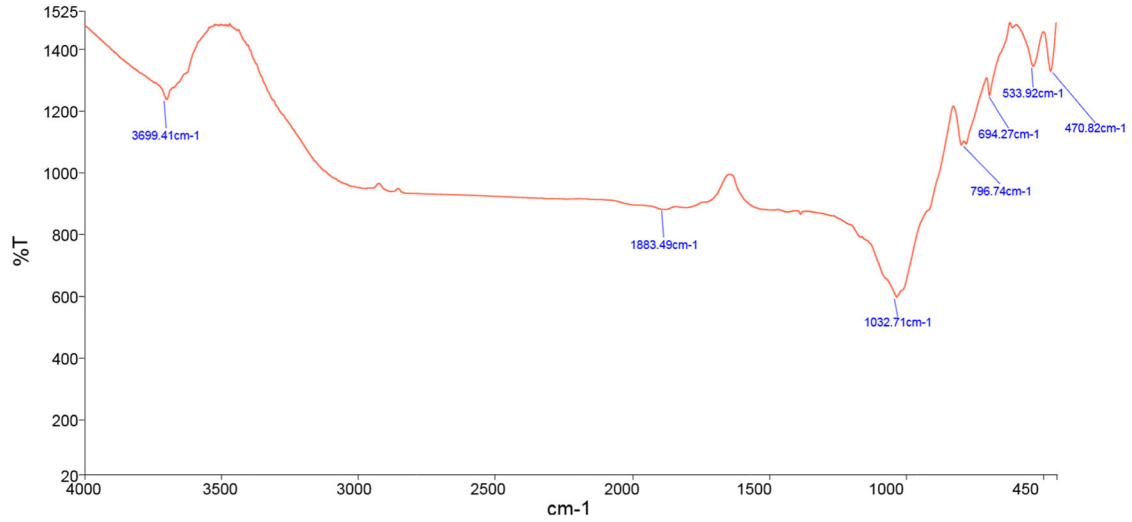


Fig. 3 FTIR spectrum of the soil collected from the study area

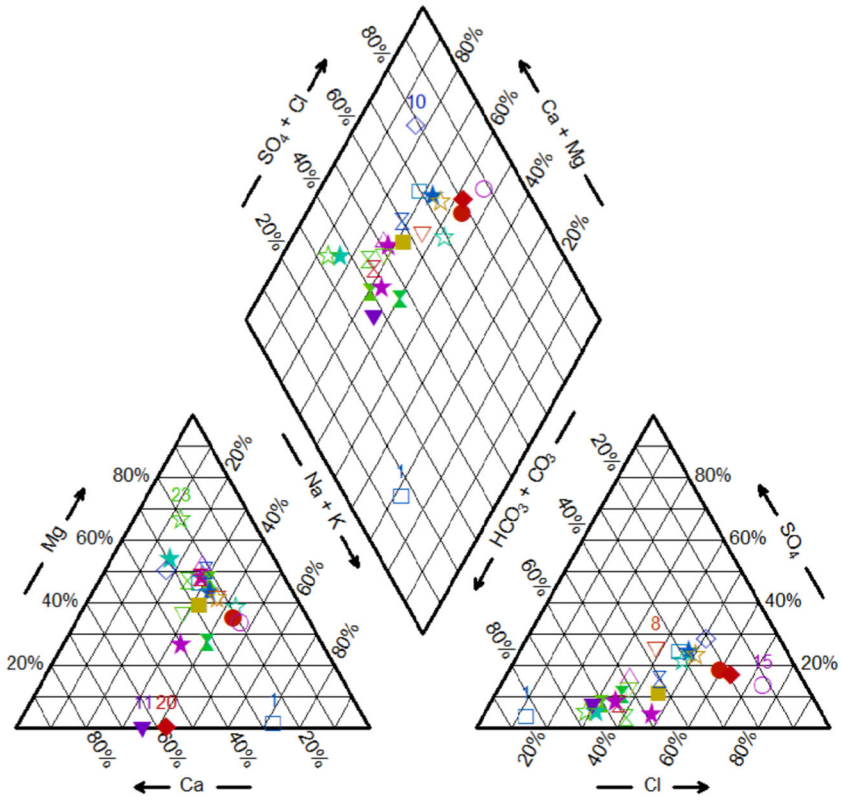


Fig. 4 Piper's diagram for geochemical classification and chemical type of groundwater of the study area

catchment of the lake to maintain the groundwater quality. On the other hand, groundwater sources located south of the lake had relatively lower concentration since an overflowing stream from Siliserh Lake might result in recharging of groundwater aquifers with fresh surface water. Based on soluble sodium concentration and its relative fraction to other cations, most of the sampled water had groundwater suitable for agricultural use. Since the water had a dominant fraction of Mg, its long-term application may result in accumulation of Mg in soil consequently increasing Mg hazard. Otherwise, the groundwater is fit for irrigation since the soil in this region is sandy–silt, which doesn't accumulate the ions. The present study concludes that the groundwater quality is significantly affected by the natural geochemical interactions but the site-associated human activities like discharge of wastewater,

intensification of agriculture and waste disposal may render the groundwater unfit for use in future.

Acknowledgements Authors acknowledge the help of Department of Irrigation, Alwar, Rajasthan to grant permission for collection of samples; the help of Dr Ram Singh, Department of Applied Chemistry, DTU for FTIR analysis; and Dr Anita Punia (JNU, Delhi) during interpretation of results is thankfully acknowledged.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Abdel Moneim AA (1988) Hydrogeology of the Nile Basin in Sohag Province. M.Sc.Thesis, Geology Department, Faculty of Science, Assiut University, Assiut, Egypt

APHA (2005) Standard methods for the examination of water and wastewater, 21st edn. American Public Health Association, Washington DC

Arveti N, Sarma MRS, Aitkenhead-Peterson JA, Sunil K (2011) Fluoride incidence in groundwater: a case study from Talupula, Andhra Pradesh. India Environ Monit Assess 172:427–443

Bishnoi SR, Brar SPS, Kumar D (1984) Underground water quality of Dhuri block, District Sangrur, Punjab. Indian J Ecol 11(2):220–228

Bureau of Indian Standards (2012) Indian standard drinking water specifications (2nd Revision) 10500. BIS, New Delhi

Central Ground Water Board (2016) Groundwater yearbook (2014–2015) Rajasthan state. Regional Office Data Centre, Western Region Jaipur

Chitrakshi, Haritash AK (2018) Hydrogeochemical characterization and suitability appraisal of groundwater around stone quarries in Mahendragarh, India. Environ Earth Sci 77:252. https://doi.org/10.1007/s12665-018-7431-5

Doneen LD (1964) Water quality for agriculture. Department of Agriculture, University of California, Davis, p 4

Eaton EM (1950) Significance of carbonate in irrigation water. Soil Sci 69:123–133

Food and Agriculture Organisation (2017) Water pollution from agriculture: a global review. FAO of United Nations, Rome

Gaillardet J, Dupre B, Louvat P, Allegre CJ (1999) Global silicate weathering and CO2 consumption rates deduced from the chemistry of large rivers. Chem Geol 59:3–30

Gibbs RJ (1970) Mechanism controlling world water chemistry. Science 17:1088–1090

Haritash AK, Kaushik CP, Kaushik A, Kansal A, Yadav AK (2008) Suitability assessment of groundwater in some villages of Rewari district in Haryana. Environ Monit Assess 145(1–3):397–406

Haritash AK, Mathur K, Singh P, Singh SK (2017) Hydrochemical characterization and suitability assessment of groundwater in Baga–Calangute stretch of Goa. India Environ Earth Sci 76:341. https://doi.org/10.1007/s12665-017-6679-5

Karaca O, Claudio C, Reddy KR (2017) Acid pond sediment and mine tailings contaminated with metals: physicochemical characterization and electrokinetic remediation. Environment Earth Sci 76 (12). https://doi.org/10.1007/s12665-017-6736-0

Kelly WP (1963) Use of saline irrigation water. Soil Sci 95:355–391

Kumar S, Lata S, Yadav J, Yadav JP (2017) Relationship between water, urine and serum fluoride and fluorosis in school-children of Jhajjar District, Haryana, India. Appl Water Sci 7:3377–3384

Kumarasamy P, Alaguraja P, Deepasaraswathi M, Sekar M, Muthuveerran P, Yuvaraj D, Thirunavukkarasu A, Arthur J, Sames R (2011) Soil suitability and water quality study in Tirupur areas using geographic information system techniques. Int J Earth Sci Eng 4:85–97

Paliwal KV (1972) Irrigation with saline water. IARI Monograph No. 2 (New Series), IARI, New Delhi, p 198

Piper AM (1953) A graphic procedure in the geochemical interpretation of water analysis. USGS Groundwater Note, No, p 12

Richards LA (1954) Diagnosis and improvement of saline-alkali soils. In: Agriculture handbook no 60. US Department of Agriculture, Washington, p 160

Sawyer CN, Mccarty PL, Parkin GF (2003) Chemistry for environmental engineering and science, 5th edn. McGraw-Hill, New York, p 752

US Geological Survey (2000) Classification of natural ponds and lakes. US Department of the Interior, US Geological Survey, Washington, DC

Wilcox LV (1955) Classification and use of irrigation water. USDA Circular 969, USDA, Washington DC, p 19

World Health Organization (2006) Guidelines for drinking water quality. First addendum to 3rd edn, vol. 1, World Health Organization, Geneva

2020 IEEE VLSI Device, Circuit and System Conference (VLSI-DCS), 18-19 July, 2020, ED MSIT SBC, Kolkata, India

Impact of Graded Back-Barrier on Linearity of Recessed Gate InAlN/GaN HEMT

Megha Sharma
Department of Applied Physics
Delhi Technological University
Delhi, India
meghasharma_phd2k18@dtu.ac.in

Rishu Chaujar
Department of Applied Physics
Delhi Technological University Delhi, India
Chaujar.rishu@dtu.ac.in

Abstract – Work presented in this paper reports the impact of graded AlGaN back-barrier on the linearity performance of lattice-matched InAlN/AlN/GaN recess gate device. Specifically, a comparison is made on the linearity of the proposed device with and without a graded back barrier layer. Various device performance matrices namely transconductance (g_m) and its higher-order derivatives, second and third-order voltage intercept point (VIP_2, VIP_3), 3rd order input intercept point (IIP_3), 3rd inter-modulation distortion (IMD_3) and 1-dB compression point have been investigated and discussed for the proposed device architecture. Results obtained from TCAD simulation of the proposed device confirms that by using graded back-barrier the linearity performance matrices of the device are improved, which further shows that the proposed device is also suitable for wireless applications.

Keywords-heterostructure, back-barrier, recess, linearity, polarization

I. INTRODUCTION

The GaN-based HEMT has been appeared as a suitable applicant for high power and high-speed applications such as military communication and navigation system. For such application commendable linearity of the device is required. The linearity of a normal single heterojunction HEMT can be improved using a graded back-barrier double heterojunction HEMT, employing InAlN barrier layer [1]and recessing the gate [2].The usually used AlGaN barrier layer has been replaced by an InAlN barrier layer.This is because when the composition Al in InAlN become 0.83 then the lattice of InAlN barrier layer matches to GaN buffer and significantly enhances the density of two-dimension electron gas(2DEG) and reliability of the device. An additional thin layer AlN is introduced between the InAlN/GaN interface in order to improve the performance of the device. AlN layer increases the conduction band discontinuity at the hetero-junction interface which further increases the concentration of 2DEG [3].In single hetero-structure, both the buffer layer and channel layer are GaN, which leads to the degradation in mobility and reliability due to the flow of 2DEG from channel to buffer. In order to enhance the 2DEG confinement, AlGaN back-barrier layer has been used, but the back-barrier layer also produces an undesirable channel at the AlGaN / GaN junction. So as to reduce the adverse effect of the unwanted channel, the graded AlGaN back-barrier used which reduces the polarization between AlGaN / GaN interface by changing the Al composition from 0 to 0.1 [4][5].A lot of work has already been done on graded back-barrier InAlN / GaN structures [4][6] which mainly focused on the normally 'ON' device operation.

In the proposed work, linearity performance of the graded back-barrier InAlN / GaN gate recessed HEMT device has been analysed. The device performance matrices have been extracted using ATLAS Silvaco TCAD device simulator [7].

II. DEVICE DESCRIPTION

Fig1 and Fig2 show the schematic view of InAlN/ AlN/GaN with AlGaN back-barrier and graded AlGaN back-barrier recess gate HEMT. The thickness of InAlN and AlN are 10nm and 1nm respectively. The unintentionally doped 2um GaN buffer layer has been composed with carrier concentration of $1x10^{16}$ cm^{-3} [8]. 14nm thick GaN channel and 25nm graded AlGaN back-barrier layers have been introduced. Both structures have a gate-source distance of 0.5 µm, gate-drain 10 µm, and gate length is 2µm respectively. The proposed architecture was passivated using SiN layer to reduce the current collapse [9]. The gate is made up of Au material and recess in the InAlN layer. Source and drain terminal are the ohmic contacts in both the structures. The parameters of GaN, AlN, InAlN, AlGaN are listed in Table1[10]–[12].The simulations are completed using the Fermi-Dirac, Shockley-Read- Hall (SRH), field-dependent mobility model. Some other physical parameters effects are taken into consideration like spontaneous polarization.

978-1-7281-1933-5/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:29:32 UTC from IEEE Xplore. Restrictions apply.

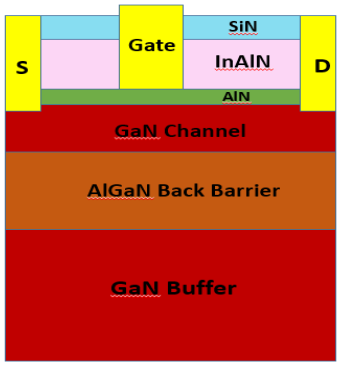


Fig.1 InAlN/AlN/GaN with AlGaN back barrier

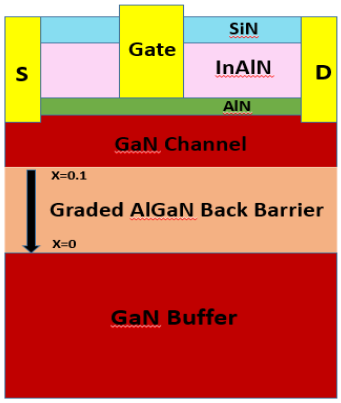


Fig 2. InAlN/AlN/GaN with Graded AlGaN back-barrier

Table 1. parameters of GaN AlN, InAlN, AlGaN

Parameter	GaN	AlN	InAlN	AlGaN
Band gap	3.42 eV	6.15 eV	4.62 eV	3.58 eV
Hole mobility	22 (cm^2 /Vs)	14 (cm^2 /Vs)	82 (cm^2 /Vs)	-
Saturated velocity	2.0 (10^7 cm/s)	2.17 (10^7 cm/s)	1.1 (10^7 cm/s)	-
Spontaneous polarization	-0.034 (C/m^2)	-0.090 (C/m^2)	-0.072 (C/m^2)	-0.038 (C/m^2)

III. Result And Discussion

Fig 3. exhibit the transfer characteristics of the proposed device and conventional device. It was found that drain current increases in graded back-barrier HEMT as compared with AlGaN back barrier. A drain voltage V_{ds}= 5V the density of maximum saturation current was obtained 0.0022A/um in an AlGaN graded back-barrier HEMT which was nearly 15.2% larger than without graded back-barrier HEMT. The higher drain current in recess gate AlGaN graded back-barrier HEMT is obtained by the reason of lower threshold voltage, in graded back-barrier HEMT, the negative shift in threshold voltage is mainly due to the reduction of negative polarization field which minimizes the confinement of carrier in the main channel.

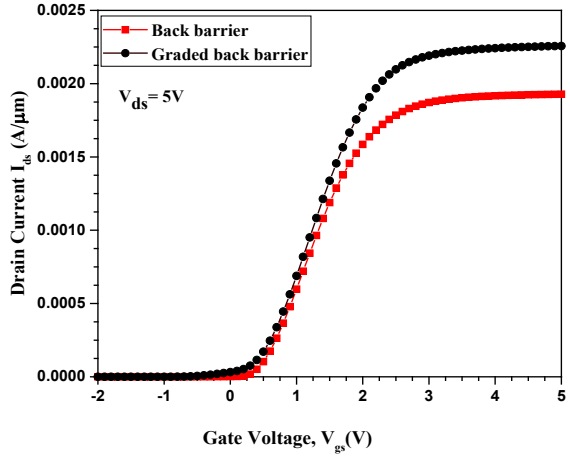


Fig 3. Id vs Vgs characteristic curve of graded and without graded back barrier

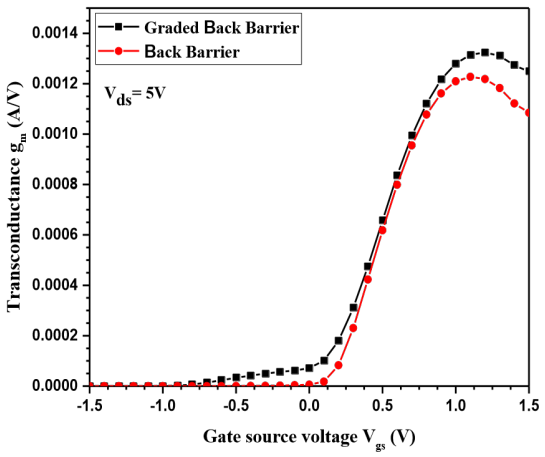


Fig 4. g_m vs V_{gs} for graded and without graded back barrier

Fig.4 presents the variation of transconductance peaks (g_m) for with graded and without graded back-barrier. It has been shown that the transconductance peak of the structure having a

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:29:32 UTC from IEEE Xplore. Restrictions apply.

graded back-barrier is more than that without grading. This increase is obtained due to the gradual decrease of Al composition, which further minimizes the effect of unwanted channel on the main channel and raises the mobility of 2DEG.

The study of non-linear behaviour essential for the power conservation of the device. The first and second-order derivative of g_m w.r.t V_{gs} is the non-linear parameters. The third-order non-linear term g_{m3} has a significant effect in wasting the useful power which affects the analog and RF performance of the device. Therefore, for high linearity of the device g_{m2} & g_{m3} should be small.

$$g_{m2} = \frac{\delta^2 I_d}{\delta V_{gs}^2} \tag{1}$$

$$g_{m3} = \frac{\delta^3 I_d}{\delta V_{gs}^3} \tag{2}$$

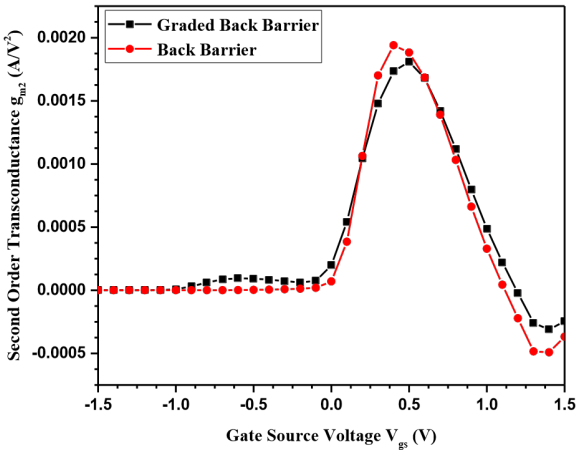


Fig.5. g_{m2} vs V_{gs} for graded and without graded back barrier

Fig 5. & 6. show the effect of graded &without graded back-barrier on the amplitude of non-linear parameters g_{m2} & g_{m3}. For graded back-barrier HEMT voltage between 0V to 0.5V g_m increases with increases in V_{gs}. The rate of increase slows down between 0.5V to 1V, hence show the downward swing in g_{m2}. As the gate voltage increase, it causes the negative value of g_{m2} because g_m decreases on increasing the gate voltage. This decrease in g_m is because at higher gate voltage two-dimensional electron gas approaches to the equilibrium value and current density become constant at higher gate voltage. From Fig 5. it has been shown that there is a decrement in peak value of g_{m2} for the graded back-barrier as result VIP2 increases which significantly improves the linearity of the device. Fig 6. shows the variation of third-order transconductance with V_{gs}. The analysis reveals that g_{m3} of graded back-barrier HEMT value much lowered as compared to without graded back-barrier HEMT.

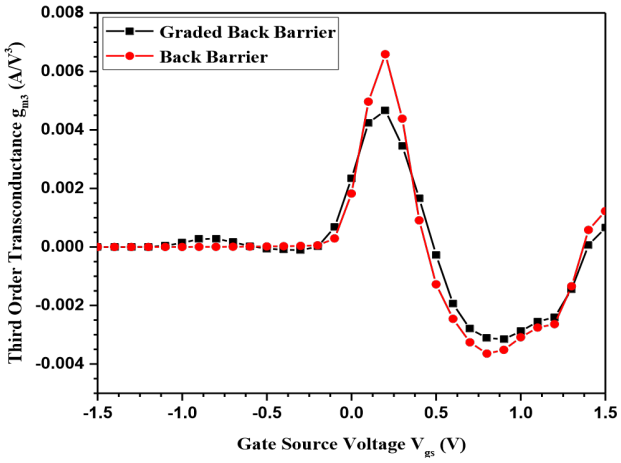


Fig. 6. g_{m3} vs V_{gs} for graded and without graded back barrier

Linearity has a significant role in all RF and microwave application hence it is essential to analyse the linearity characteristics of the device. There are some linearity parameters VIP2, VIP3, IIP3, IMD3 & 1-dB compression point as define below [13].

$$VIP2 = 4 * \frac{g_m}{g_{m2}} \tag{3}$$

$$VIP3 = \sqrt{24 * \frac{g_m}{g_{m3}}} \tag{4}$$

$$IIP3 = 4 * \frac{g_m}{g_{m3} * R_s} \tag{5}$$

$$IMD3 = \frac{9}{2} * (VIP)^3 * g_{m3} \tag{6}$$

$$1 - dB\ compression\ point = 0.22 \sqrt{\frac{g_m}{g_{m3}}} \tag{7}$$

Where VIP2 & VIP3 are the second and third-order voltage intercept point where IIP3 is the input intercept. For analog & microwave applicationVIP2,VIP3, IIP3 & 1-dB compression point should be high and IMD3 should be as small as possible.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:29:32 UTC from IEEE Xplore. Restrictions apply.

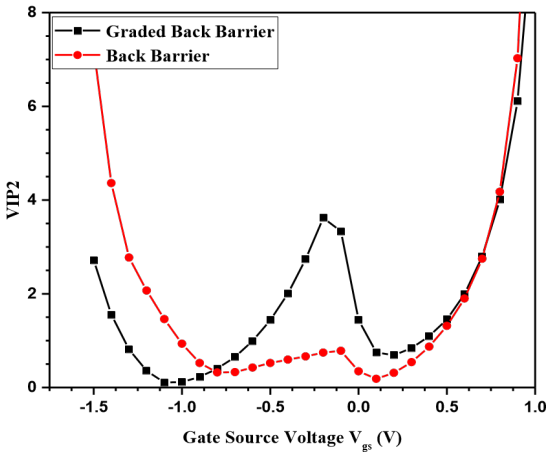


Fig.7. VIP$_2$ vs V$_{gs}$ for graded and without graded back barrier

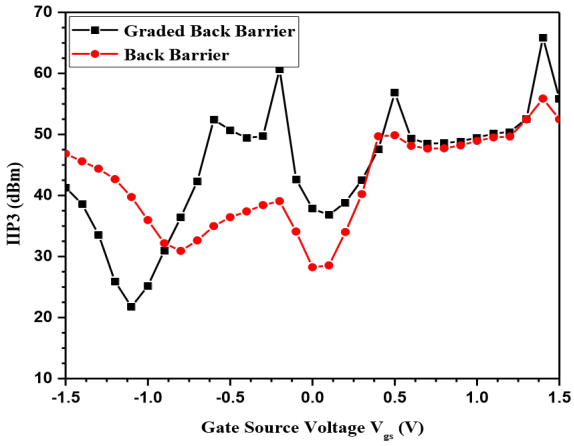


Fig. 9. IIP$_3$ vs V$_{gs}$ for graded and without graded back barrier

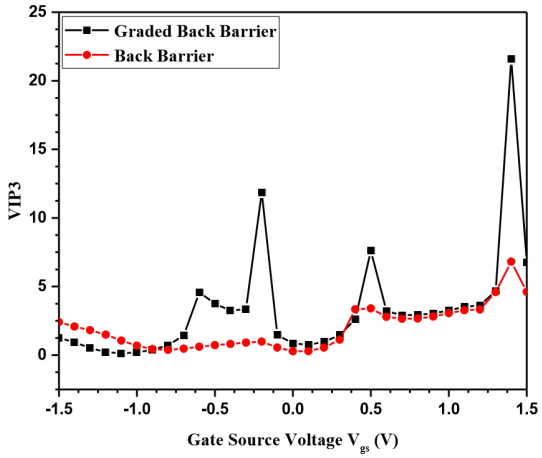


Fig .8. VIP$_3$ vs V$_{gs}$ for graded and without graded back barrier

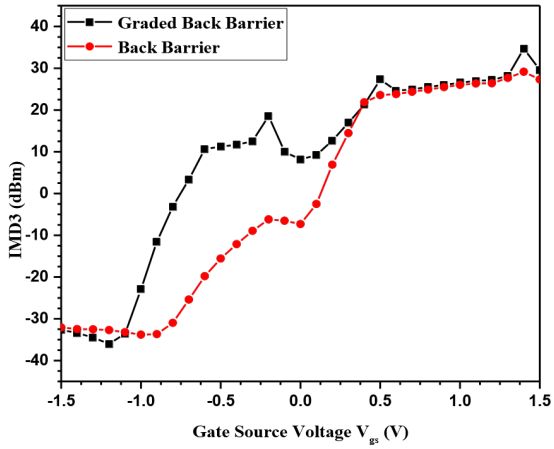


Fig. 10. IMD3 vs V$_{gs}$ for graded and without graded back barrier

Fig.7 and Fig8. exhibit the VIP2 & VIP3 linearity parameters of graded and without graded back barrier. Since VIP$_2$ and VIP$_3$ have an inverse proportionality relationship with g_{m2} & g_{m3} respectively, hence lower the value of g_{m2} and g_{m3} better will be the linearity performance. It has also been verified that the highest value of VIP2 and VIP3 are obtained towards the lower gate bias with graded back-barrier, which proves the superior linearity performance of the graded back-barrier device.

Fig 9. & 10. exhibit the curve of IIP3 and IMD3 as a function of gate bias. It has been observed that at lower voltage, higher peak of IIP3 obtained for graded back barrier hence lower biasing is required to reduce the distortion. Fig 10 shows that at higher gate biasing lower peak of IMD3 obtained for graded back barrier.

Fig 11 Showing that 1 dB compression point for graded AlGaN back-barrier is 80.11% higher which emphasizes the superiority in linearity performance of graded AlGaN back-barrier for operation in microwave circuit.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:29:32 UTC from IEEE Xplore. Restrictions apply.

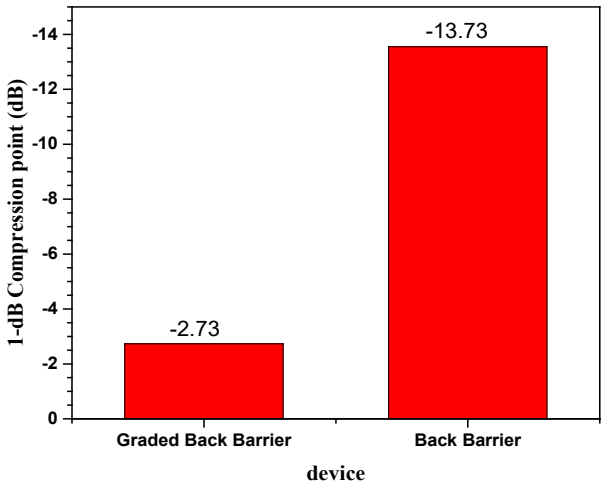


Fig. 11. 1 dB compression point of graded and without graded back barrier

.

IV. Conclusion

We have discussed the linearity performance of InAlN/ GaN Recess Gate HEMT with the graded back barrier. It is observed that variation in Al composition in back-barrier plays an important role in improving the linearity performance of the proposed architecture. Various device performance matrices obtained using exhausted device simulation clearly indicates that by using the graded back-barrier, linearity performance of the device is improved. It was also noticed that the higher value of VIP2 and VIP3 are obtained for a graded back-barrier device. In addition to this, it can also be concluded that IIP3 and IMD3 point of the device improves with the graded back-barrier concept.

REFERENCES

[1] M. Neuburger, T. Zimmermann, and E. Kohn, "UNSTRAINED InAlN / GaN HEMT STRUCTURE," vol. 14, no. 3, pp. 785–790, 2004.

[2] R. Wang *et al.*, "Gate-recessed enhancement-mode InAlN/AlN/GaN HEMTs with 1.9-A/mm drain current density and 800-ms/mm transconductance," *IEEE Electron Device Lett.*, vol. 31, no. 12, pp. 1383–1385, 2010, doi: 10.1109/LED.2010.2072771.

[3] Y. L. and E. Y. C. Niraj Man Shrestha1, "Simulation study on electrical characteristic of AlGaN/GaN high electron mobility transistors with AlN spacer layer," *Jpn. J. Appl. Phys.*, vol. 53, 2014.

[4] P. A. B. Layer *et al.*, "electronics Theoretical Study of InAlN / GaN High Electron Mobility Transistor (HEMT) with a," pp. 0–9, 2019.

[5] C. Yao *et al.*, "High-performance AlGaN-based solar-blind avalanche photodiodes with dual-periodic III-nitride distributed Bragg reflectors," *Appl. Phys. Express*, vol. 10, no. 3, 2017, doi: 10.7567/APEX.10.034302.

[6] T. Han, H. Zhao, X. Peng, and Y. Li, "Control of short-channel effects in InAlN/GaN high-electron mobility transistors using graded AlGaN buffer," *Superlattices Microstruct.*, vol. 116, pp. 207–214, 2018, doi: 10.1016/j.spmi.2018.02.031.

[7] F. G. Della Corte, F. Pezzimenti, S. Bellone, and R. Nipoti, "Numerical simulations of a 4H-SiC BMFET power transistor with normally-off characteristics," *Mater. Sci. Forum*, vol. 679–680, pp. 621–624, 2011, doi: 10.4028/www.scientific.net/MSF.679-680.621.

[8] S. Faramehr, K. Kalna, and P. Igic, "Modeling of 2DEG and 2DHG in i-GaN capped AlGaN/AlN/GaN HEMTs," *Proc. Int. Conf. Microelectron. ICM*, no. Miel, pp. 81–84, 2014, doi: 10.1109/MIEL.2014.6842090.

[9] Z. H. Feng, Y. G. Zhou, S. J. Cai, and K. M. Lau, "Enhanced thermal stability of the two-dimensional electron gas in GaN/AlGaN/GaN heterostructures by Si3N4 surface-passivation-induced strain solidification," *Appl. Phys. Lett.*, vol. 85, no. 22, pp. 5248–5250, 2004, doi: 10.1063/1.1828231.

[10] H. Morkoç, R. Cingolani, and B. Gil, "Polarization effects in nitride semiconductor device structures and performance of modulation doped field effect transistors," *Solid. State. Electron.*, vol. 43, no. 10, pp. 1909–1927, 1999, doi: 10.1016/S0038-1101(99)00146-X.

[11] O. Ambacher *et al.*, "Two-dimensional electron gases induced by spontaneous and piezoelectric polarization charges in N- And Ga-face AlGaN/GaN heterostructures," *J. Appl. Phys.*, vol. 85, no. 6, pp. 3222–3233, 1999, doi: 10.1063/1.369664.

[12] O. Ambacher *et al.*, "Pyroelectric properties of Al (In) GaN / GaN hetero- and," *J. Phys. Condens. Matter*, vol. 14, pp. 3399–3434, 2002.

[13] S. P. Kumar, A. Agrawal, R. Chaujar, R. S. Gupta, and M. Gupta, "Device linearity and intermodulation distortion comparison of dual material gate and conventional AlGaN/GaN high electron mobility transistor," *Microelectron. Reliab.*, vol. 51, no. 3, pp. 587–596, 2011, doi: 10.1016/j.microrel.2010.09.033.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:29:32 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Second International Conference on Inventive Research in Computing Applications (ICIRCA-2020)
IEEE Xplore Part Number: CFP20N67-ART; ISBN: 978-1-7281-5374-2

Impact of Metal Strip on Nanoscale Double Gate Overlap Tunnel FET

Anushka Jain
Department of Applied Physics
Delhi Technological University
Delhi-42, India
anushkajain.dtu@gmail.com

Saransh Sharma
Department of Applied Physics
Delhi Technological University
Delhi-42, India
saransh.sharma249@gmail.com

Rishu Chaujar
Department of Applied Physics
Delhi Technological University
Delhi-42, India
chaujar.rishu@dtu.ac.in

***Abstract*— In this paper, a gradual advancement in the development of high performance TCAD Simulation of Tunnel Field Effect Transistor (TFET) for better switching application is introduced by increasing Ion current. This is achieved by implanting a metal strip in each oxide layer near the gate and source junction in a Double Gate Tunnel FET. Furthermore, the use of HfO_2 instead of SiO_2 as the oxide layer has been examined. The combination of metal strips in the oxide layer of HfO_2 offers a higher I_{ON}/I_{OFF} ratio, steep subthreshold slope (SS) and significant change in the performance of transconductance (g_m), transconductance efficiency (TGF) and gate-to-drain capacitance (C_{gd}). Hence, at low gate voltages the proposed device could be a better substitute for ultra-low power applications. All simulated results presented here are carried out with the aid of 2-D Silvaco ATLAS Software.**

Keywords*— *Band-to-band tunnelling, High-K dielectric, Subthreshold slope, Transconductance, Transconductance Efficiency, Silvaco Atlas Software

I. INTRODUCTION

As technology is growing, the market is shifting towards miniature size, but with the demand of similar functionality in the scaled down versions of the original. Thus, power consumption and size of the device are the main constraints. To meet these requirements, MOS devices were heavily scaled down which resulted in short-scaling effects and thus, engineered an advanced electronic device called as Tunnel Field Effect Transistor (TFET) [1]. Since TFET's subthreshold swing (SS) is not restricted to 60mV/dec with lower off current and threshold voltage while being immune to short-channel effects, has made it apt for low-power applications and has attracted many researcher's interests.

TFET's structure is almost closer to MOSFET but with different switching mechanisms. MOSFET's mechanism works on modulating thermionic emission over a barrier while in TFET, switching mechanism is done by modulating a quantum tunnel through a barrier which provides greater operating speed. Basis of TFET's physics is the phenomenon of non-localized band-to-band tunneling at the junctions of varying doping i.e. at source-channel and channel-drain for which a similar algorithm is used for its simulation [10].

Gate positioning has been one of the very few important modifications and has attracted researcher's interest. From various literature [2], it is evident that the modification varies from where the gate covers the area near the source to the model where it completely overlaps with the source. This paper has chosen the one, where the gate and source overlap completely. This yields better I_{ON} current. To provide greater control of the gate over the electrostatics of the channel, the use of Double Gate Tunnel FET (DG-TFET) has been inspected.

This paper investigates the combination of DG-TFET with its gate by overlapping the source called as double gate overlap tunnel FET (DGO-TFET). A comparison is done between this new and advanced device with its conventional counterpart i.e. TFET. Leakage of current at gate is caused when gate voltage is high enough to increase the field of electrons in vertical direction [11]. Thus, for a device to work in lower gate voltage is desirable. Now, to accomplish a greater subthreshold slope (SS) and I_{ON}/I_{OFF} ratio, an additional introduction of metal strips in the oxide layer of SiO_2 at the source-channel junction has been looked into.

This paper analyzes this evolved device, Metal-Strip Double-Gate Overlap Tunnel FET (MS DGO TFET). To further achieve a higher ratio of On-Off current, material with high-k dielectric, such as HfO_2 (k = 21) has been considered instead of SiO_2 This change in oxide material has been evaluated and compared in terms of their performance parameters such as transconductance (g_m) and gate-to-drain capacitance (C_{gd}). The rest of the paper is compiled as follows.

Section II describes the structure and simulation setup of the introduced device and compares it with its conventional device and its variant. It also describes the simulation setup and methodology used to generate electric field profiles. Section III compares the Id-Vg curve and discuss the performance of other parameters like switching ratio (I_{ON}/I_{OFF}), subthreshold slope (SS), transconductance (g_m), transconductance efficiency (TGF) and gate-to-drain capacitance (C_{gd}). The concluding statements are specified in the final Section.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:25:46 UTC from IEEE Xplore. Restrictions apply.

II. DEVICE STRUCTURE, DESCRIPTION AND SIMULATION SETUP

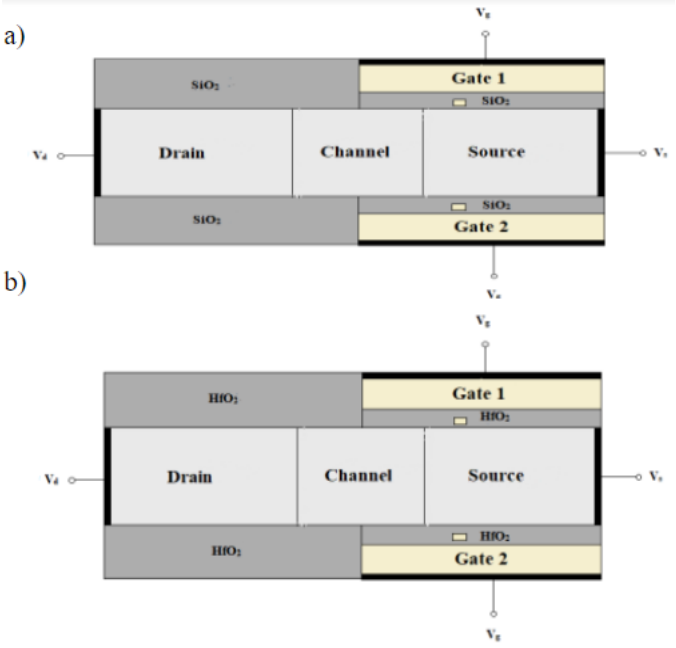


Fig.1. Schematic cross-sectional view of proposed TFET with (a) SiO_2 as oxide and (b) HfO_2 as oxide.

Fig.1 shows the cross-sectional view of a conventional GOTFET with a proposed metal strip implant. Table 1 contains proposed structure's dimensions and values of certain parameters. As shown in Fig.1., ultra-low I_{off} at low V_{GS} values is achieved by reverse biased p-i-n structure, while the gate overlapping the source region results in higher I_{ON} at high V_{GS}. Proper material used as the gate contacts enhance I_{on} and the desired V_t can be obtained. The channel length of the device is 30 nm, the gate metal thickness is 1.5 nm and the oxide thickness is also 1.5 nm. With decrease in gate oxide thickness, the drain current I_d increases, however, as intel's process reported 1.2 nm as the minimum thickness of gate oxide to avoid gate tunneling, i.e. a gate oxide of 1.5 nm thickness is used. Gate dielectric made up of 1.5 nm thick HfO_2 (relative permittivity □ =22) layer has been used, to make it reliable and to avoid tunneling. The drain is n+ type doped ($5\times10^{18}cm^{-3}$), body is p+ type doped (1×10^{17} cm^{-3}) and source is p+ type doped (1×10^{20} cm^{-3}). The abrupt change in doping levels of source and drain leads to suppression of ambipolarity effect [3].

The subthreshold slope (SS=51.5) and I_{on}/I_{OFF} ratio approximately to the 10^{10} are achieved by the metal strips implanted in the oxide layer. The metal strip is located near the junction of source and channel, in the oxide layer to enhance the performance of the device. Both metal strips and gate are made up of aluminium.

A. Simulation Methodology

Atlas device simulator is used to perform all the simulations [4]. The models used during simulations are as follows: non-local BTBT, band gap narrowing (BGN), Shockley-Read-Hall (SRH) for carrier recombination, concentration and field dependent mobility model. Instead of local tunneling methods such as Kane model, the non-local BTBT model is considered. The non-local BTBT model is used as it takes into account the dynamic change in electric field at each point in the tunneling path, so it is dependent on the tunneling path's band profile [3]. In the tunneling path of the local tunneling model constant electric field is assumed at each point as it uses triangular barrier approximation [5]. The quantum mesh is defined at various regions to calculate tunneling carriers. To keep numerical efficiency in mind, the mesh is carefully refined in the tunneling regions and is coarse elsewhere. Beside this, Newton method is used for numerical solution.

TABLE I: PARAMETERS OF THE PROPOSED STRUCTURE IN THIS PAPER

Parameter	Value	Unit
Source doping	1×10^{20}	$/cm^3$
Channel doping	1×10^{17}	$/cm^3$
Drain doping	5×10^{18}	$/cm^3$
Source length	100	nm
Channel length	30	nm
Drain length	100	nm
Metal strip thickness	0.5	nm
Metal strip length	5	nm
Gate work function	4.5	eV

B. Fabrication Feasibility

Fabrication is the major limitation of such device. The main concern regarding obtaining this device can be production of metal layer inside gate oxide layer. This can be easily solved by using atomic layer deposition (ALD) [6]. ALD passes through vapour phase to deposit thin films. Thickness can be controlled at 0.1 nm level using ALD. The temperature required for this technique is less than 3500 °C which is quite modest. Our device can be easily obtained by ALD, as the metal layer is 0.5 nm thick and 5 nm long [10]. To obtain the required thin film deposition, there is a specific temperature window, known as "ALD temperature window".

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:25:46 UTC from IEEE Xplore. Restrictions apply.

If temperature is kept outside this window, then it leads to slower reactions and power growth rates. ALD is also used for deposition of Au, Pt, Hf, Cu. Chemical vapor deposition can also be used for depositing thin films [7].

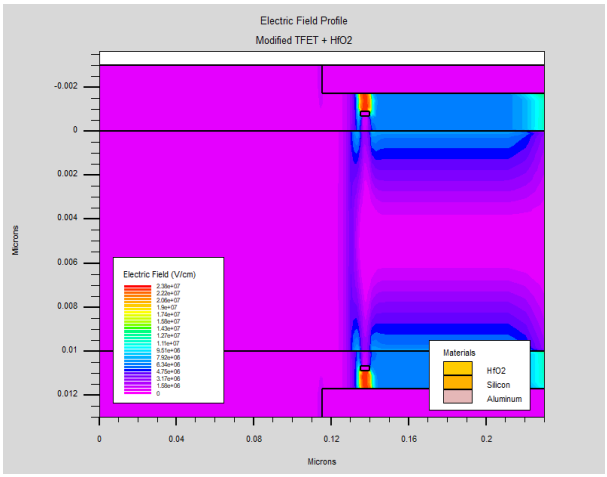
III. RESULTS AND DISCUSSIONS

This section will discuss and analyze DC and AC characteristics of the three devices and compare them in order to achieve a more advanced version of TFET with better DC and AC Characteristics.

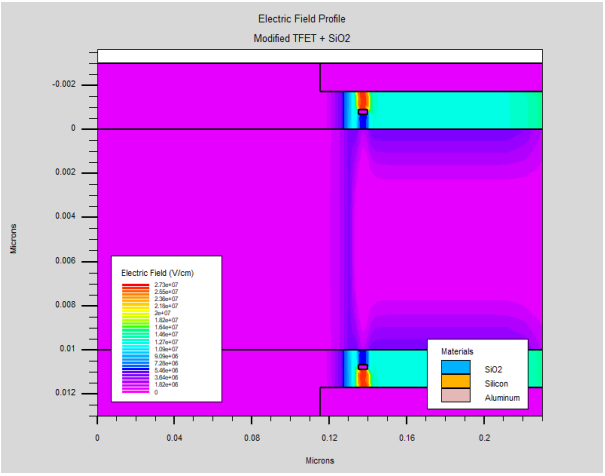
C. DC Characteristics

The section analyses the DC Characteristics such as the electric field profiles and transfer characteristics i.e. Id/Vg curve of all the three devices and compares them.

a)



b)



c)

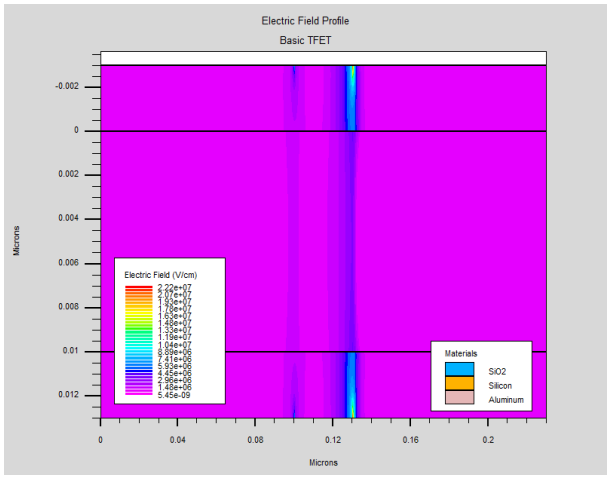


Fig.2. Comparison of Electric Field Profiles of a)MS-DGO-TFET with HfO2 as oxide, b)MS-DGO-TFET with SiO2 as oxide and c) Tunnel FET

Fig.2. shows the electric field profile for proposed structures and conventional TFET. It can be observed that introduction of metal strips leads to an increased electric field at the interface of channel and source. Consequently, high electric field causes greater accumulation of electrons, which in turn leads to more energy banding under the metal strip [4]. This also leads to better SS and maximum I_{ON}/I_{OFF} ratio.

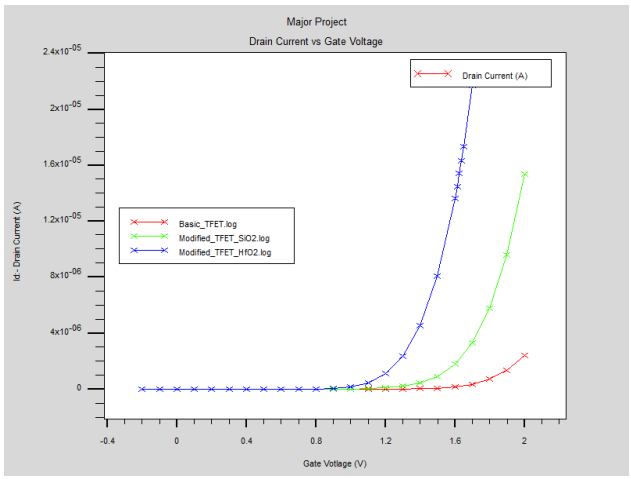


Fig. 3. Represents the Transfer Characteristics Curve of Id/Vg for Vdrain=3V of MS-DGO-TFET with HfO2 as oxide, MS-DGO-TFET with SiO2 as oxide and Tunnel FET.

MS-DGO-TFET with SiO_2 displays greater current than TFET as the presence of double gate provides more electrostatic control over the channel in which drain field lines cannot affect or disturb the channel. The device, MS-DGO-TFET with HfO_2 shows maximum on-state current. From this it can be inferred that the overlapping of gates and implantations of metal strips has been taken as advantage. Replacing HfO_2 with SiO_2 has led to further increase in the current. Drain current versus gate voltage graph explains how MS-DGO-TFET with HfO_2 is able to reach higher drain current for small voltages as compared to MS-DGO-TFET with SiO_2, which in turn performs even better than the conventional Tunnel FET.

Switching ratio is given as I_{ON}/I_{OFF}. Greater value of switching ratio is desirable. It is the figure of merit of having high performance. These values were calculated for vdrain=1V as shown in table 2. It is evident from the table that MS-DGO-TFET with HfO_2 as oxide has maximum Ion current and greater switching ratio than the conventional device of TFET.

Another key factor is the Subthreshold Slope (SS) which is the inverse of the subthreshold swing of the device. From [12-13], subthreshold swing of the device is defined as, "The change in gate voltage which must be applied in order to create a one decade increase in the output current". There are two types of subthreshold slope. First is the *point subthreshold slope* which is defined at a particular value of V_{gs}. While the second type is the *average subthreshold slope*, which is

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:25:46 UTC from IEEE Xplore. Restrictions apply.

defined over a range of V_{gs} values. To achieve this, the proposed work has used the point subthreshold slope.

Subthreshold Slope (SS) depends on C_d which increases with Na. Reducing capacitance of depletion layer (C_d) is not possible due to bulk Si technology. SS also depends on gate-oxide capacitance (C_{ox}), which increases with decreasing oxide thickness (T_{ox}).

D. AC/Analog Parameters

A relative analysis has been performed in the section between MS-DGO-TFET with HfO2 as oxide (Modified_HfO2), MS-DGO-TFET with SiO2 as oxide (Modified_SiO2) and the conventional counterpart, TFET, in terms of transconductance (gm), transconductance efficiency (TGF) and gate-to-drain capacitance (Cgd).

For simulation of these parameters, "AC FREQ=1.0e+6" was added after voltage sweep. Small signal AC simulation was performed for each gate bias for 1MHz frequency.

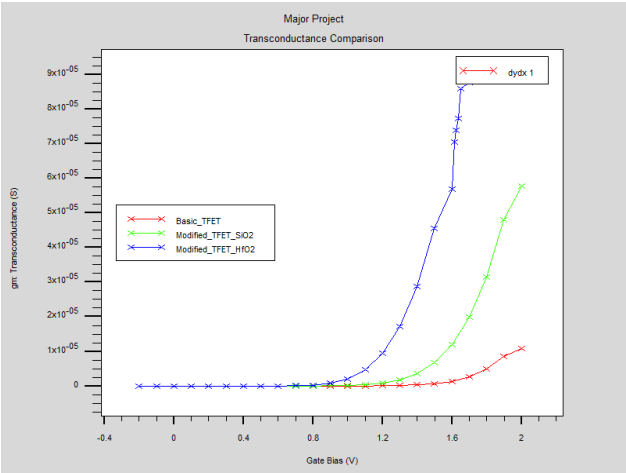


Fig. 4. Represents the overlay graph of Transconductance (gm) as calculated from Id/Vg curve and plotted against gate voltage.

The comparison in the correlation of transconductance and gate voltage of all three devices is shown in Fig. 4. Transconductance (g_m) is an expression of the performance of the semiconductor device. According to [14], transconductance (g_m) is defined as "the ratio of the change in drain current to the change in gate voltage over a defined, arbitrarily small interval on the drain-current v/s gate-voltage curve".

$$g_m = dI_d/dV_{gs} \qquad for\ V_{ds}=constant$$

In general, the larger its value, the greater is its capability of delivering the gain (amplification), when all the other factors remain unchanged [15]. From figure 4, it is evident that the transconductance of MS-DGO-TFET with HfO2 as oxide (Modified_HfO2), shows a rapid increase because of the increased tunneling current. This positive influence is due to the fact that work function of metal strip is more effective at lower gate voltages when compared to the conventional TFET [6]. Another reason for increased performance is HfO_2 as oxide instead of SiO_2. Thus, its implantation in DGO-TFET increases g_m by many folds.

Transistors are efficient generators of transconductance in subthreshold region [8] [16-17]. For analog integrated circuits, the supreme parameter in subthreshold region is the ratio of transconductance to drain current [9]. This ratio is coined as transconductance generation efficiency. Therefore, it is desirable to have a greater transconductance generation efficiency.

TABLE II: COMPARISON BETWEEN PROPOSED DEVICES AND CONVENTIONAL DEVICE FOR V_{DRAIN}=1V.

Parameters Name. Symbol & Units.	**Modified HfO2**	**Modified SiO2**	**TFET**
On - State Current Ion (A/m)	5.99E-06	3.45E-06	7.53E-07
Off - State Current Ioff (A/m)	2.00E-16	8.36E-18	1.61E-16
Switching Ratio I_{on}/I_{off}	3.00E+10	4.12E+11	4.67E+09
Subthreshold Slope SS (mV/dec)	51.5	56.45	60.00
Transconductance Efficiency *TGF*	44.70%	40.70%	35.29%

Transconductance Efficiency (TGF) is given as:

$$TGF = gm/Id = ln(10)/SS \quad (SS:Subthreshold\ Slope)$$

This figure of merit quantifies how much gm will be obtained for the amount of bias current invested. It is also clear from the equation that, high dielectric materials exhibit higher TGF as compared to lower dielectric structures. As the dielectric permittivity increases the gm/Id ratio decreases and becomes worse at k=30.

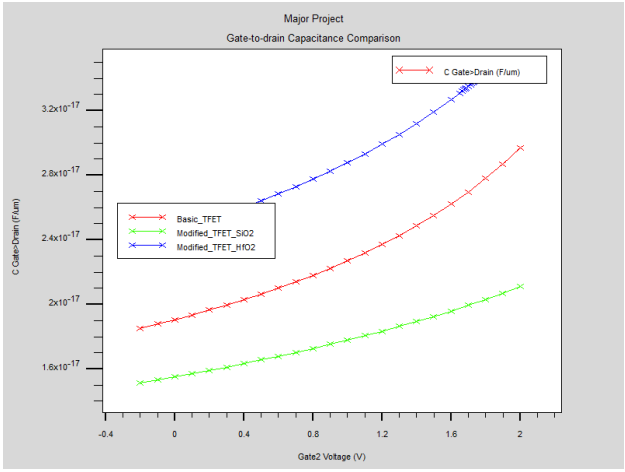


Fig. 5. Represents an overlay graph of Gate-to-drain capacitance (C_{gd}) plotted against gate voltage.

Fig. 5 shows that the extracted C_{gd} is a function of gate voltage (V_{gs}) at 1Mhz. For total gate capacitance (C_{gg}), C_{gd} alone is accountable as C_{gs} decreases exponentially with the increase in V_{gs}. This is due to the existence of potential barrier

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:25:46 UTC from IEEE Xplore. Restrictions apply.

at the source end. Value of C_{gd} enhances with the increasing value of V_{gs} owing to lowering in potential barrier at the drain end [18]. This C_{gd} is also known as Miller Capacitance and it should be as low as possible. The proposed devices have lower values of C_{gd}, in the order of 10^{-17} which is desirable.

IV. CONCLUSION

This paper has proposed a high performing simulation using TCAD of double gate overlapped TFET with the source and metal strips implanted in its oxide layer at source-channel junction. The software has a lot of potential and with right commands and features, an accurate, practical and calculative guess can be made. After getting practical outputs, it can be fabricated and can be used industrially. The device performance was analyzed and compared with its conventional counterpart and a modified version where the oxide layer is replaced by HfO_2. The results presented in this paper indicate that the modified device with HfO_2 as oxide has greater switching ratio (3.00E+10) than TFET (4.67E+09) and a lower subthreshold slope (SS) of 51.5 mV/dec than the TFET (60.0mV/dec). As the subthreshold slope decreases, transconductance efficiency has been increased. Thus, the proposed device has enhanced characteristics than its conventional counterpart and can be explored further for application in low-power. More exhaustive research can be done on this proposed device either with change in materials, position of metal strip, number of metal strip or addition of another modification.

ACKNOWLEDGMENT

The authors of this paper would like to express their gratitude to the Department of Applied Physics, Delhi Technological University for their exceptional and continuous guidance and support.

REFERENCES

[1] J Madan, RS Gupta, R Chaujar, Analytical drain current formulation for gate dielectric engineered dual material gate-gate all around-tunneling field effect transistor, Japanese Journal of Applied Physics 54 (9), 094202, 2015

[2] Kumar, S., Goel, E., Singh, K., Singh, B., Singh, P. K., Baral, K., et al. (2017). 2-Danalytical modeling of the electrical characteristics of dual-material double-gate TFETs with a SiO2/HfO2stacked gate-oxide structure. IEEE Transactions on Electron Devices, 64(3), 960.

[3] J. Madan and R. Chaujar, "Temperature Associated Reliability Issues of Heterogeneous Gate Dielectric—Gate All Around—Tunnel FET," in IEEE Transactions on Nanotechnology, vol. 17, no. 1, pp. 41-48, Jan. 2018, doi: 10.1109/TNANO.2017.2650209.

[4] A. U. s. Manual, "Silvaco," Santa Clara, CA, 2010.

[5] E. Vittoz and J. Fellrath, "CMOS analog integrated circuits based on weak inversion operation," IEEE J. Solid-State Circuits, vol. 12, no. 3, pp. 224–231, Jun. 1977.

[6] Chandan, B.V., Gautami, M., Nigam, K. et al. Impact of a metal-strip on a polarity-based electrically doped TFET for improvement of DC and analog/RF performance. J Comput Electron 18, 76–82(2019). https://doi.org/10.1007/s10825-018-1280-z

[7] Johnson, R.W., Hultqvist, A., Bent, S.F.: A brief review of atomiclayer deposition: from fundamentals to applications. Mater. Today 17, 236–246 (2015)

[8] Guo, L., Leobandung, E., Zhuang, L., Chou, S.Y.: Fabrication and Characterization of Room Temperature Silicon Single Electron Memory. American Vacuum Society, New York (1997)

[9] F. Silveira, D. Flandre, and P. G. A. Jespers, "A gm/ID based method for the design of CMOS analog circuits and its application to the synthesis of a silicon-on-insulator micropower OTA," IEEE J. Solid-State Circuits, vol. 31, no. 9, pp. 1314–1319, Sep. 1996.

[10] C. Shen, L. T. Yang, E. -. Toh, C. -. Heng, G. S. Samudra and Y. -. Yeo, "A new robust non-local algorithm for band-to-band tunneling simulation and its application to Tunnel-FET," 2009 International Symposium on VLSI Technology, Systems, and Applications, Hsinchu, 2009, pp. 113-114.

[11] S.K. Kurinec, in Encyclopedia of Materials: Science and Technology, 2001.

[12] www.Silvaco.com

[13] Physics of Semiconductor Devices, S. M. Sze. New York: Wiley, 3rd ed., with Kwok K. Ng, 2007, chapter 6.2.4, p. 315, ISBN 978-0-471-14323-9.

[14] "IEEE-ICDCS conference proceeding", 2012 International Conference of Devices Circuits and Systems (ICDCS), 03/2012

[15] Radio Frequency Transistors: Principles and practical applications,EDN Series for Design Engineers, Helge Granberg. ISBN 0080571433, 9780080571430

[16] Leonardo Barnobi, Mariana Siniscalchi, Berardi Sensale-Rodriguez. "TFET-Based Circuit Design Using the Transconductance Generation Efficiency ${g}_{m}/{I}_{d}$ Method", IEEE Journal of the Electron Devices Society, 2015.

[17] B. Senale-Rodríguez et al., "Perspectives of TFETs for low power analog ICs," 2012 IEEE Subthreshold Microelectronics Conference (SubVT), Waltham, MA, 2012, pp. 1-3, doi: 10.1109/SubVT.2012.6404307.

[18] K. Sivasankaran, P S Mallick, N. Murali, Kiran Kumar. "RF and stability performance of Double Gate Tunnel FET", 2012 International Conference on Emerging Electronics, 2012.

[19] Kumar, N. M. (2019). ENERGY AND POWER EFFICIENT SYSTEM ON CHIP WITH NANOSHEET FET. Journal of Electronics, 1(01), 52-59.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:25:46 UTC from IEEE Xplore. Restrictions apply.

Improving Software Maintainability Predictions using Data Oversampling and Hybridized Techniques

Ruchika Malhotra
Discipline of Software Engineering, Department of Computer Science and Engineering
Delhi Technological University, Delhi, India
ruchikamalhotra2004@yahoo.com

Kusum Lata
Department of Computer Science and Engineering,
Delhi Technological University, Delhi, India
kusumlata@dtu.ac.in

***Abstract*— Software systems being developed in today's era are inherently large and complex. Maintaining these software is a big challenge before the software industry. Since software maintenance demands a high cost, this activity becomes even more challenging. In order to reduce the maintenance cost, it becomes crucial to know the maintainability of software modules/ classes in the initial stages of software development. Many efforts have been made to identify the maintainability of modules in the initial development stages in which prediction models have a lot of roles. Prediction models are trained from past historical data and should consist of an adequate number of instances of low maintainability and high maintainability class/modules. But this is not usually the case, and because of this, we are not able to train the prediction models properly. This situation shows data imbalance. In this direction, in this paper, we will handle the imbalanced data problem so that prediction models can be properly trained. We apply four oversampling techniques in this paper and train the prediction models for maintainability by hybridized techniques. The results of the paper advocate the effectiveness of examined oversampling techniques along with the hybridized classification techniques to develop competent maintainability prediction models.**

Keywords—Software Maintenance, Imbalanced data, Hybridized techniques.

I. INTRODUCTION

Maintenance is an essential dimension of software quality. Maintainability means how easily a software module or class can be modified [1]. Since the size and complexity of software systems are increasing due to technological progress, it becomes very important to develop maintainable software. It is also necessary to develop software in a maintainable way because maintenance charges a good cost and effort of software development [2]. But in reality, the maintainability of a software component can be determined in the actual maintenance phase. At that stage, it would be very problematic to control the cost of maintenance. The cost of maintenance can be controlled if we come to know in the initial development stages, how maintainable a particular module be? In the literature, researchers have made many efforts to forecast software maintainability in the initial phases of software development, and through their efforts, many maintainability prediction models have been developed [3-7]. Software metrics that are the measurable factors have been used as the independent variables to develop models to predict software maintainability with the application of different statistical and ML techniques [8-10]. The software maintainability prediction models are trained from past historical data comprising of data points. Each data point consists of software metrics capturing different characteristics of class/module and maintainability.

Maintainability of a software class/module will be low if more revisions in the lines of source code are needed to modify that module in the maintenance phase and vice versa [11]. It is very important that we have adequate of instances of low maintainability and high maintainability class in order to train any predictive model efficiently. In the real scenario, the low maintainability classes are far less in number as compare to high maintainable classes. This means we have less number of data points representing one class (low maintainability) and more number of data points representing another (high maintainability) class. If this is the situation, we will call it a data imbalance problem. In the data imbalance situation, the class with a small number of instances is called a minority class, and the class whose instances are more is called the majority class. In context to imbalanced data problem in this paper, low maintainability class is referred to as minority class, and high maintainability class is the majority class. in this paper, accurate prediction of low maintainability class is our prime concern as they increase the cost of maintenance.

Many competent machine learning (ML) and statistical techniques fail to predict the data points of each class accurately when the data is imbalanced [12]. Realizing how important it is to handle the data imbalance problem, in this paper, we apply four oversampling techniques. The examined oversampling techniques help to balance data by increasing the minority class data points. Also, we use hybridized classification techniques to construct prediction models for software maintainability. As we know that these days search-based techniques are becoming quite popular. These techniques search for a vast solution space and find an optimal solution [13-14]. For developing prediction models for maintainability, we use hybrid classification techniques in this study. These techniques associate search-based and ML techniques as one single technique to be used for a prediction task [15]. Studies in the past have examined the remarkable capability of hybrid classification techniques in software quality predictive modeling tasks [7][16-17].

978-1-7281-6929-3/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 10:21:25 UTC from IEEE Xplore. Restrictions apply.

However, very few studies in the past have applied these techniques for software maintainability prediction.

Therefore, the prime contributions of this study are:

(i) Effective pre-processing the imbalanced data by oversampling before the development of models

(ii) To construct prediction models for predicting the low maintainability classes with the aid of hybrid classification techniques after data oversampling. We address the following research questions.

RQ#1) When the data is imbalanced, what is the predictive performance of software maintainability prediction models? Do oversampling techniques improve the performance of software maintainability prediction models?

RQ#2) Which is the best oversampling to develop competent software maintainability prediction models? Is the best oversampling techniques, statistically better than other examined oversampling techniques?

We use three application packages of Apache open-source software, namely Apache Betwixt, Apache IO, and Apache Ivy, in this study for our experiment. We validate the training and testing of models by ten-fold cross-validation and assess the performance of constructed models with the help of GMean and Balance metrics.

The layout of the various section of the study is depicted as follows: Section II is designated for related work. Section III explains the research methodology. Section IV is dedicated to the experimental setup. We present results and their analysis in section V. In section VI, we document threats to validity and finally conclude the paper in section VII.

II. Related Work

There are a lot of studies in the literature wherein the software maintainability prediction is made by validating the software metrics quantify different characteristics of the software. The data imbalance problem is not addressed in any study as these studies predict the maintainability in the form of a continuous variable (number of changes made to code in the maintenance period)[3-7]. The aim of the present study is to predict low maintainability classes efficiently by addressing the data imbalance problem. The data imbalance problem has been dealt with in software quality prediction for software defect prediction and change prediction by various researchers.

The study by Pelayo and Dick [18] used synthetic minority oversampling technique (SMOTE) to develop competent software defect prediction from imbalanced defect datasets. Siers and Islam [19] employed a cost-sensitive classification for defect-prediction in the imbalanced scenario. The study by Zheng [20] proposed a cost-sensitive boosting technique to boost the performance of neural networks trained from imbalanced defect data. Tan et al. [21] pre-processed imbalanced datasets by using different data resampling techniques and constructed competent change-prediction models. A study by Malhotra and Kamal [22] applied various data oversampling methods to improve the performance of defect prediction models. Bashir et al. [23] pre-processed the imbalanced defect prediction datasets by taking care of borderline and noisy datapoints. It is to be noted that all of the above studies addressed the issue of imbalanced data while predicting defective module/classes.

For constructing software quality prediction modes, many studies have applied hybridized classification techniques. In the literature, there are some studies [7][16-17] that ascertain the applicability of hybridized classification techniques for developing software quality prediction modes. According to a literature review conducted by Malhotra and Khanna [22] on 61 studies in the field of software quality domain, hybrid techniques have the remarkable predictive capability. Still, very few studies have used these techniques for software maintainability prediction. It inspires us to use these techniques to construct maintainability prediction models. So in this study, we first balance the imbalanced datasets by using oversampling techniques namely: Synthetic Minority Oversampling Technique: SMOTE [25], Borderline Synthetic Minority Oversampling Technique: BSMOTE [26], Safe-Level Synthetic Minority Oversampling Technique: SSMOTE [27], and Adaptive Synthetic Oversampling technique: Adasyn [28]. We then develop maintainability prediction models by four hybridized techniques namely Fuzzy Learning based on Genetic Programming Grammar Operators and Simulated Annealing (GFS-SP) [29], Fuzzy Learning based on Genetic Programming Grammar Operators (GFS-GPG) [29], Fuzzy Learning based on Genetic Programming (GFS-GP) [29], and Fuzzy rule approach based on genetic cooperative-competitive learning (GFS-GCCL)[30].

III. Research Methodology

In this section, the procedure of data collection, software metrics used as predictor variables to construct software maintainability models, and performance metrics to evaluate the performance of models is briefly explained.

A. Data Collection Procedure

We have used three application packages of Apache, namely Apache Betwixt, Apache Io, and Apache Ivy, in this study. We have collected datasets corresponding to these application packages with the Defect Collection and Reporting (DCRS) tool [31]. Two consecutive versions of the software fed to the DCRS tool as input. In the output, we get as many data points as there are the common classes in the analyzed versions. Each such datapoint consist of eighteen software metrics, number of lines of code added (*LSoC_added*) from the previous version to next version, number of lines of code deleted (*LSoC_deleted*), number of lines of code modified (*LSoC_modified*), and the total number of lines of code changes (*LSoC_total*). We then transform *LSoC_total* variable into the Maintainability variable after discretizing. The variable Maintainability would be a binary variable with two values: Low-maintainability (LOWM) and High-maintainability (HIGHM). Here, LOWM means that a particular class will demand more revisions in its lines of code as compared to HIGHM class. In Table I, we enlist all the three software used in this paper including the name of the software, its version, number of common classes in the versions analyzed (NCC), number of low maintainability classes (NLM), Number of high maintainability classes (NHM), and imbalance ratio (IBR) where IMR is (NHM/NLM).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 10:21:25 UTC from IEEE Xplore. Restrictions apply.

B. Software Metrics used as Predictors

Eighteen software metrics qualifying different characteristics are used as predictors in this study.

Table I. Software project details

Software	Version	NCC	NLM	NHM	IBR
Apache Betwixt	0.6 – 0.7	245	27	218	08.07
Apache IO	2.0.1 – 2.2	248	10	238	23.80
Apache Ivy	1.4.1 – 2.2.0	371	28	343	13.25

These metrics are reported in Table II. Due to space constraint, we have not given the detailed definition of each metric. The detailed definition of metrics described in Table II can be referred from [32]. The response or dependent variable is binary variable: Maintainability that consists of two values, LOWM and HIGHM.

C. Assessment in Software Maintainability Prediction

For a binary classification problem like in this paper, the performance of models is judged with the confusion matrix given in Table III. The confusion matrix contains four quantities: True Positive (t-pos), False Positive (f-pos), True Negative (t-neg), and False Negative (f-neg). The values of various performance measures like True Positive Rate (TPRate), False Positive Rate (FPRate), True Negative Rate (TNRate), GMean, and Balance, etc. can be computed from the values given in confusion matrix. We use GMean and Balance performance measures in this study to record the performance of software maintainability prediction models as the use of these metrics is advocated for the imbalanced data problem [12]. Further, in this study, Positive class is corresponding to LOWM class, and the Negative class corresponds to HIGHM class.

GMean is the geometric mean of TPRate and TNRate. Balance measures the Euclidian distance between the pair of TPRate and FPRate and the optimal value of TPRate and FPRate.

$$TPRate = \frac{t-pos}{t-pos+f-neg} \quad \text{(i)}$$

$$NRate = \frac{t-neg}{t-neg+f-pos} \quad \text{(ii)}$$

$$GMean = \sqrt{TPRate * TNRate} \quad \text{(iii)}$$

$$Balance = 1 - \sqrt{\frac{(0-FPRate)^2 + (1-TPRate)^2}{2}} \quad \text{(iv)}$$

$$where\ FPRate = \frac{f-pos}{f-pos+ t-neg} \quad \text{(v)}$$

IV. Experimental Setup

This segment of the study provides information regarding the experimental setup. As discussed in section III, we have used three datasets extracted from Apache application packages details of which are provided in Table I. As the datasets are highly imbalanced as depicted from the values of IBR in Table I, we pre-process the datasets with the help of four oversampling techniques. The description of the oversampling techniques given ahead in this section. After data oversampling, for developing the software maintainability prediction models, we use hybridized classification techniques described in the last part of this section. During the course of training the models, we use a ten-fold cross-validation approach that divides the training dataset into ten partitions, out of which nine partitions are used for training the model while on the remaining one partition model is tested. We assess the performance of developed models with the performance measures discussed in section III and perform statistical analysis with the Friedman test [33] and the Wilcoxon Signed-Rank test [33].

A. Techniques used for data oversampling

To deal with imbalanced data, we have used four oversampling techniques whose concise description is as follows:

(i) Synthetic Minority Oversampling Technique (SMOTE): This technique generates artificial or synthetic data points of minority class along the direction line joining minority class datapoint and its k-nearer neighbours.

Table II. Software metrics used

Software Characteristic	Metric	Description
Size	WMC	Weighted methods per class
	NPM	Number of public methods
	LOC	Lines of source code
	DAM	Data access metric
	AMC	Average method complexity
Coupling	CBO	Coupling between objects
	RFC	Response for class
	Ca	Afferent coupling
	Ce	Efferent coupling
	IC	Inheritance coupling
	CBM	Coupling between the class methods
Cohesion	LCOM, LCOM3	Lack of cohesion among methods
Inheritance	CAM	Cohesion among methods of class
	NOC	Number of children
	DIT	Depth of inheritance
	MFC	Measure of functional abstraction
Composition	MOA	Measure of aggregation

Table III. The Confusion Matrix

	Predicted Positive	Predicted Negative
Actual Positive	t-pos	f-neg
Actual Negative	f-pos	t-neg

(ii) Borderline-Synthetic Minority Oversampling technique (BSMOTE): As an extension of SMOTE, in this technique, only borderline minority class datapoints are increased. If *k*-nearer neighbors class datapoints of a minority belong to majority class, that minority class datapoint is said to be borderline.

(iii) Safe-Level Synthetic Minority Oversampling technique (SSMOTE): This technique is another enhancement of basic SMOTE. Before performing oversampling of minority class

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 10:21:25 UTC from IEEE Xplore. Restrictions apply.

datapoint, its safe-level is determined. For a minority class datapoint, D, safe-level of D would be equal to the number of minority class instances in the *k*-nearer neighbors of D.

(iv) Adaptive Synthetic Oversampling technique (Adasyn): Unlike SMOTE and its variants, in this technique oversampling of minority class is done in accordance with its degree of difficulty in learning. A difficult to learn datapoint is oversampled by creating more number of data points corresponding to it and vice versa.

B. Hybridized Classification Techniques.

The techniques used in this study for learning training software maintainability prediction models are fuzzy rule-based classifications techniques, whose brief description is as follows: GFS-GP and GFS-GCCL techniques use genetic programming (GP) for classification with fuzzy rules. GP starts with a population of randomly generated solutions and iteratively converts a population of solutions into a new generation of the population by applying genetic operations. GFS-GPG uses grammar-based GP for discovering effective fuzzy rules. GFS-SP uses a simulated annealing based search in order to search for an effective solution in fuzzy rule bases.

V. Results and Analysis

A. RQ#1) When the data is imbalanced, what is the predictive performance of software maintainability prediction models? Do oversampling techniques improve the performance of software maintainability prediction models?

In order to answer this research question, we develop software maintainability prediction models using hybridized techniques in two steps (i) by using original imbalanced datasets (ii) after applying oversampling techniques to imbalanced datasets. In both of the scenarios, we assess the performance of the developed models with respect to GMean and Balance.

Table IV. GMean results of Software Maintainability Prediction Models

	GFS-GCCL	GFS-GP	GFS-GPG	GFS-SP	Mean
Apache Betwixt					
Adasyn	71.18	66.50	69.60	68.41	**68.92**
BSMOTE	66.40	56.53	60.73	57.51	**60.29**
SSMOTE	70.67	65.95	66.50	68.63	**67.94**
SMOTE	71.68	61.33	64.28	68.72	**66.50**
No-resampling	0.00	19.25	19.20	18.93	14.34
Apache IO					
Adasyn	74.81	79.28	72.56	78.39	**76.26**
BSMOTE	73.79	68.60	53.84	68.75	**66.25**
SSMOTE	78.00	79.28	77.27	76.50	**77.76**
SMOTE	77.35	78.00	79.07	66.10	**75.13**
No-resampling	0.00	54.66	54.54	54.31	40.88
Apache Ivy					
Adasyn	69.58	85.59	68.49	64.56	**72.06**
BSMOTE	69.19	65.85	63.10	66.27	**66.10**
SSMOTE	61.33	61.36	68.91	66.08	**64.42**
SMOTE	65.85	67.49	68.86	65.31	**66.88**
No-resampling	26.73	26.73	26.71	26.65	26.70

The performance of the models after employing oversampling techniques along with No-resampling is recorded in Table IV and Table V, respectively, for GMean and Balance. In Tables IV and V, No-resampling refers to the case when Software maintainability prediction models are developed using imbalanced datasets. As shown in Table IV and Table V, in the No-resampling situation, the performance of the models is very poor. The average GMean for Apache Betwixt, IO, Ivy was 14.34, 40.88, and 26.70, respectively. Similarly, average Balance values were reported to be merely 31.25, 45.20, and 34.34 for Apache Betwixt, IO, and Ivy, respectively. For Apache Betwixt dataset, GMean values ranged from 56.53 to 71.68 after oversampling. For Apache IO and Ivy datasets, GMean values ranged from 53.84 to 79.28 and 61.33 to 85.59, respectively, as shown in Table IV. Similarly, as reported in Table V, For Apache Betwixt dataset, Balance values ranged from 55.41 to 71.31 after oversampling. For Apache IO and Ivy datasets, Balance values ranged from 50.45 to 79.27 and 60.54 to 82.25, respectively. After oversampling, for Apache Betwixt dataset, mean GMean values ranged from 60.29 to 68.92, and the Adasyn technique has reported the highest GMean of 68.92. For Apache IO, the range of average GMean values was 66.25 to 77.76 after applying oversampling techniques, and SSMOTE has the highest average of GMean of 77.76, and in case of Apache Ivy, the range of average GMean values was 64.42 to 72.06 with Adasyn of highest average GMean of 72.06 as shown in Table IV. Similarly, on analyzing Table V, we observe very much improvement in the performance of software maintainability prediction models with respect Balance performance metric.

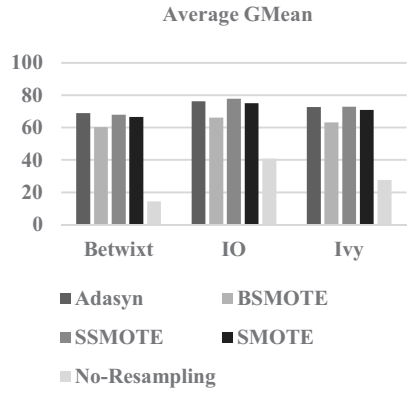


Fig.1 Average GMean of all datasets after oversampling.

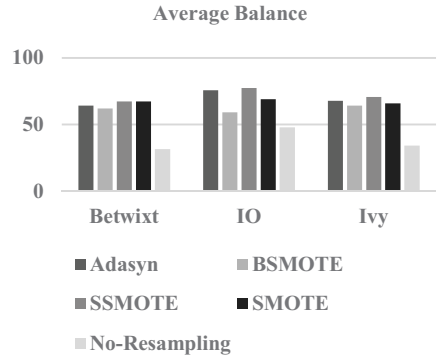


Fig.2 Average Balance of all datasets after oversampling.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 10:21:25 UTC from IEEE Xplore. Restrictions apply.

For Apache Betwixt dataset, average Balance values ranged from 59..15 to 68.68 and Adasyn technique has reported highest average Balance of 68.68 For Apache IO, the of was average Balance values ranged from 62.56 to 77.36 and SSMOTE has highest average Balance of 77.76 and in case of Apache Ivy the range of average Balance was 64.19 to 71.04 with Adasyn of highest average Balance of 71.04, as shown in Table V. Fig.1 and Fig.2 show the average GMean and average Balance performance of all three datasets after applying oversampling along with No-resampling. Therefore, it is clearly evident from Fig.1 and Fig.2 that investigated oversampling techniques improve the predictive performance of the software maintainability prediction models.

The analysis the results of the developed software maintainability prediction models before and after oversampling, it is evident that the oversampling of the datasets improved the performance of the models.

Table V. Balance results of Software Maintainability Prediction Models

	GFS-GCCL	GFS-GP	GFS-GPG	GFS-SP	Mean
Apache Betwixt					
Adasyn	70.77	66.40	69.16	68.38	**68.68**
BSMOTE	65.35	55.41	59.78	56.04	**59.15**
SSMOTE	70.56	65.53	66.40	67.41	**67.48**
SMOTE	71.31	61.20	64.07	68.49	**66.27**
No-resampling	29.29	31.91	31.91	31.87	31.25
Apache IO					
Adasyn	72.34	79.27	72.48	76.48	**75.14**
BSMOTE	70.97	64.4	50.45	64.43	**62.56**
SSMOTE	77.94	79.27	76.38	75.83	**77.36**
SMOTE	77.25	77.71	78.84	63.54	**74.33**
No-resampling	29.29	50.5	50.5	50.49	45.20
Apache Ivy					
Adasyn	69.55	82.25	67.90	64.47	**71.04**
BSMOTE	69.18	65.65	62.88	66.25	**65.99**
Safe-SMOTE	60.54	61.27	68.90	66.06	**64.19**
SMOTE	65.65	67.40	68.66	65.16	**66.72**
No-resampling	34.34	34.34	34.34	34.34	34.34

Table VI. Friedman Test Results

Oversampling Techniques	Rank with respect to GMean	Rank with respect to Balance
Adasyn	Rank 1 (3.96)	Rank 1 (4.04)
SSMOTE	Rank 2 (3.79)	Rank 2 (3.71)
SMOTE	Rank 3 (3.67)	Rank 3 (3.67)
BSMOTE	Rank 4 (2.5)	Rank 4 (2.50)
No Resampling	Rank 5 (1.08)	Rank 5 (1.08)
p-value	0.00	0.00

B. RQ#2) Which is the best oversampling technique to develop competent software maintainability prediction models? Is the best oversampling techniques, statistically better than other examined oversampling techniques?

In order to answer this research question, we analyze the performance of each oversampling technique on the software maintainability prediction models by two non-parametric statistical tests, Friedman Test and Wilcoxon Signed Rank Test on the GMean and Balance results at a significance level of $\alpha = 0.05$. First, we statistically analyze the performance of prediction models with the aid of the Friedman test that assigns a mean rank to each technique according to the performance of each oversampling technique in terms of GMean and Balance. The results of the Friedman test are documented in Table VI for GMean and Balance. As shown in Table VI, the results of the Friedman test for investigated oversampling are significant both for GMean and Balance, p-value < 0.05. Adasyn technique has obtained the highest rank with respect to GMean and Balance. The second best rank is obtained by SSMOTE technique. The results for the Friedman test indicate that No Resampling situation has obtained the worst rank both for GMean and Balance measures. To investigate the second part of this RQ, we further carry our investigation with the help of the Wilcoxon Signed-Rank test with Bonferroni correction (= 0.05/4=0.0125). We make a pairwise comparison for the best oversampling (Adasyn) technique with other oversampling techniques used in the study. The results of the Wilcoxon Signed-Rank test are recorded in Table VII.

Table VII. Wilcoxon Test Results for Oversampling Techniques

Resampling Technique Pair	p-value : GMean	p-value : Balance
Adasyn-SSMOTE	0.477 (Not Sig+)	0.657(Not Sig+)
Adasyn-SMOTE	0.308 (Not Sig+	0.347(Not Sig+)
Adasyn-BSMOTE	0.005 (Sig+)	0.005(Sig+)

As shown in Table VII, the Adasyn technique is statistically different (Sig+) from BSMOTE (p-value < 0.0125), both with respect to GMean and Balance. However, from Table VII, it can be seen that the Adasyn technique is not statistically different (Not Sig+) in terms of its performance from SSMOTE and SMOTE.

On statistically analysing on the performance of the software maintainability prediction models we observed that Adasyn, SSMOTE, and SMOTE are equally competent oversampling techniques for developing effective models to predict low maintainability classes accurately.

VI. THREATS TO VALIDITY

We have developed software maintainability prediction models in this study by pre-processing imbalanced datasets with the help if oversampling techniques and evaluated the performance of the developed modes with respect to robust and stable evaluators, GMean and Balance. Therefore, there is no internal validity threat. Also, we don't have the threat to construct validity in this study as the software metrics that are used to construct the prediction models have been widely validated in the software engineering domain. We have reported the results after rigorous statistical analysis performed with the non-parametric statistical test. Therefore, the threat to conclusion validity not exists in this study. The datasets used in this study are extracted from the application packages of Apache written in Java programming language. So, there may be the threat that our result can vary for software developed in other programming languages.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 10:21:25 UTC from IEEE Xplore. Restrictions apply.

VII. CONCLUSIONS OF THE STUDY

Developing the competent prediction models to correctly predict the low maintainability classes is the prime objective of this study. Because the early prediction of such classes help to enhance the software quality by rigorous testing such classes. However, the imbalanced data poses difficulty to train a model for this task. This motivates us to develop competent models by pre-processing the imbalanced datasets.

With this objective, this paper examines the effect of oversampling techniques for developing prediction models to predict software maintainability. We conducted experiments on three datasets extracted from different Apache application packages, four oversampling techniques, and four hybridized classification techniques. The predictive of software maintainability prediction models that use No resampling (imbalanced datasets) were compared with the models that use oversampling methods with respect to GMean and Balance. The results are statistically validated with the help of the Friedman test and Wilcoxon Signed Rank test. The results of our experiment support that oversampling increases the predictive performance of the software maintainability prediction models. The Adaysn, SSMOTE, and SMOTE techniques proved to be the best oversampling techniques to balance the data distribution and develop competent software maintainability prediction models. In order to enhance the generalizability of our results, in future we plan to broaden the present investigation by collecting more diverse datasets. In this study, we investigated the data level solution to imbalanced data problem. We plan to undertake a few algorithmic level solutions to the imbalanced data problem for software maintainability prediction. We also plan the extend our work for inter-project and inter-release validation.

REFERENCES

[1] IEEE Std. 610.12-1990. Standard Glossary of Software Engineering Terminology, IEEE Computer Society Press, Los Alamitos, CA, 1993.

[2] I. Sommerville, Software Engineering, Addison-Wesley, 1995.

[3] M. O. Elish, H. Aljamaan and I. Ahmad, I. "Three empirical studies on predicting software maintainability using ensemble methods," *Soft Computing*, vol. 19, no. 9, 2015, pp. 511-2524.

[4] Y. Zhou and H. Leung, "Predicting object-oriented software maintainability using multivariate adaptive regression splines," *Journal of Systems and Software*, vol. 80, no. 8, 2007, pp. 1349-1361.

[5] T. S. Quah and & M. M. T. Thwin, "Application of neural networks for software quality prediction using object-oriented metrics, " *In Proceeding of IEEE International Conference on Software Maintenance,* ICSM 2003, pp. 116-125.

[6] H. Alsolai and M. Roper, "Application of Ensemble Techniques in Predicting Object-Oriented Software Maintainability," *In Proceedings of the Evaluation and Assessment on Software Engineering,* ACM*,* 2019, pp. 370-373.

[7] L. Kumar and S.K. Rath, "Software maintainability prediction using hybrid neural network and fuzzy logic approach with parallel computing concept," *International Journal of System Assurance Engineering and Management*, vol. 8, no. 2, pp. 1487-1502.

[8] A. Kaur and K. Kaur, K. "Statistical comparison of modelling methods for software maintainability prediction," *International Journal of Software Engineering and Knowledge Engineering*, vol. 23*,* no. 6, 2013, pp. 743-774.

[9] W. Zhang, L. Huang, V. Ng, and J. Ge, "SMPLearner: learning to predict software maintainability". *Automated Software Engineering*, vol. *22,* no. 1, 2015,pp. 111-141

[10] C. Van Koten and A. Gray, "An application of bayesian network for predicting object-oriented software maintainability," *Information and Software Technology*, vol. 48, no. 1, 2006, pp. 59-67.

[11] J. Al Dallal, "Object-oriented class maintainability prediction using internal quality attributes," *Information and Software Technology*, vol. *55,* no. 11,2013, pp. 2028-2048.

[12] H. He, E. A. Garcia, "Learning from imbalanced data. *IEEE Transactions on knowledge and data engineering*," vol. 21*,* no.9, 2009, pp. 1263-1284.

[13] J. Clarke, J. J. Dolado, M. Harman, R. Hierons, B. Jones, M. Lumkin, , ... and M. Shepperd, "Reformulating software engineering as a search problem," *IEE Proceedings-software*, vol. 150, no. 3, pp. 161-175.

[14] M. Harman, "The relationship between search based software engineering and predictive modeling," In *Proceedings of the 6th International Conference on Predictive Models in Software Engineering*, ACM, 2010, p. 1.

[15] R. Malhotra, "Search based techniques for software fault prediction: current trends and future directions," In *Proceedings of the 7th International Workshop on Search-Based Software Testing,* ACM, 2014, pp. 35-36.

[16] D. Rodríguez, R. Ruiz, J. C. Riquelme and J. S. Aguilar–Ruiz, "Searching for rules to detect defective modules: A subgroup discovery approach, "*Information Sciences*, 2012, pp. 14-30.

[17] L. L. Minku and X. Yao, "Software effort estimation as a multiobjective learning problem," *ACM Transactions on Software Engineering and Methodology (TOSEM)*, vol. 22, no. 4, 2013, 35.

[18] L. Pelayo, S. Dick, "Applying novel resampling strategies to software defect prediction" *In IEEE 2007 Annual Meeting of the North American Fuzzy Information Processing Society*, NAFIPS 2017, pp. 69-72.

[19] M. J. Siers and M. Z. Islam, M. Z. "Software defect prediction using a cost sensitive decision forest and voting, and a potential solution to the class imbalance problem," *Information Systems*, vol. 51, 2015, 62-71.

[20] J. Zheng, "Cost-sensitive boosting neural networks for software defect prediction," *Expert Systems with Applications*, vol. *37,* no. 6, 2010, pp. 4537-4543.

[21] M. Tan, L. Tan, S. Dara and C. Mayeux, "Online defect prediction for imbalanced data," In *2015 IEEE/ACM 37th IEEE International Conference on Software Engineering*, vol. 2, 2015, pp. 99-108. IEEE.

[22] R. Malhotra and S. Kamal, "An empirical study to investigate oversampling methods for improving software defect prediction using imbalanced data." *Neurocomputing*, vol. 343, pp. 120-140.

[23] K. Bashir, T. Li, C. W. Yohannese and M. Yahaya, SMOTEFRIS-INFFC: Handling the challenge of borderline and noisy examples in imbalanced learning for software defect prediction. *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 1, 2020, pp. 917-933.

[24] R. Malhotra, M. Khanna and R. R. Raje, "On the application of search-based techniques for software engineering predictive modeling: A systematic review and future directions," *Swarm and Evolutionary Computation*, vol. 32, 2017, 85-109.

[25] N. V. Chawla, K.W. Bowyer, L. O. Hall and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique*," Journal of artificial intelligence research,* vol. 16, 2002, pp. 321-357.

[26] H. Han, W. Y. Wang and B. H. Mao, "Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning," In *International conference on intelligent computing*. Springer, Berlin, Heidelberg, August 2005, pp. 878-887.

[27] C. Bunkhumpornpat, K. Sinapiromsaran and C. Lursinsap, "Safe-level-smote: Safe-level-synthetic minority over-sampling technique for handling the class imbalanced problem," *In Pacific-Asia conference on knowledge discovery and data mining*, Springer, Berlin, Heidelberg, April 2009, pp. 475-482.

[28] H. He, Y. Bai, E. A. Garcia and S. Li, "ADASYN: Adaptive synthetic sampling approach for imbalanced learning," In *2008 IEEE*

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 10:21:25 UTC from IEEE Xplore. Restrictions apply.

International Joint Conference on Neural Networks, June 2008, pp. 1322-1328.

[29] L. Sánchez, I. Couso and J. A. Corrales, "Combining GP operators with SA search to evolve fuzzy rule based classifiers," *Information Sciences*, vol. 136 no.4, 2001,pp. 175-191.

[30] H. Ishibuchi, T. Nakashima and T. Murata, T. "Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 29, no. 5, 1999, pp. 601-618.

[31] R. Malhotra, N. Pritam, K. Nagpal and P. Upmanyu, (2014, September). "Defect collection and reporting system for git based open source software," *In IEEE International Conference on Data Mining and Intelligent Computing*, ICDMIC September 2014, pp. 1-7.

[32] http://gromit.iiar.pwr.wroc.pl/p_inf/ ckjm/metric.html

[33] D. W. Zimmerman and B. D. Zumbo, "Relative power of the Wilcoxon test, the Friedman test, and repeated-measures ANOVA on ranks. *The Journal of Experimental Education*, vol. 62, no. 1, 1993, pp. 75-86.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 10:21:25 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Second International Conference on Inventive Research in Computing Applications (ICIRCA-2020)
IEEE Xplore Part Number: CFP20N67-ART; ISBN: 978-1-7281-5374-2

Inpainting of Irregular Holes in a Manuscript using UNet and Partial Convolution

Amreen Kaur, Ankit Raj, N. Jayanthi, S.Indu
Department of Electronics and Communication Engineering
Delhi Technological University (Formerly Delhi College of Engineering), Delhi 110042, India
amreenrocks.98@gmail.com (Amreen Kaur), rajankit606@gmail.com (Ankit Raj),
njayanthi@dce.ac.in (N. Jayanthi), s.indu@dce.ac.in(S.Indu)

***Abstract* - Image inpainting of irregular holes in the old handwritten manuscripts is a challenging task. It requires finding the missing part in the manuscript and then filling it with an appropriate pixel. It needs to fill it in such a way that words or letters in that manuscript can be recovered. There are many degraded manuscripts available in museums and libraries all across the globe. This degradation can be due to human mishandling or certain environmental factors (fire, heat, water, microorganisms, pests, and other vermin). Even if it is tired very hard to protect our written heritage from all this, since it is written on natural materials (wood, paper), it is prone to inherent vice. A novel method is presented to inpaint the irregular holes in manuscripts with appropriate pixels. Deep learning, which tends to solve many complex problems like image classification, object detection, image segmentation, helps in achieving image inpainting as well. UNet is proposed to generate the mask of an irregular hole part in a manuscript image and partial convolution, which takes two input masks and the image for inpainting.**

Keywords—Image inpainting, UNet, deep learning, partial convolution, manuscript.

I. INTRODUCTION

Manuscripts are the witness of the past. From the moment human beings figured out the need to keep a note of things, manuscripts came into existence. Manuscripts that have existed in antiquity for centuries have considerable significance. Famous personalities in the past used to write manuscripts to keep a record of their thoughts, learnings, and incidents that happened around them. These literary sources help historians and archaeologists to draw a clear picture of the ancient times. Basically, the timeline leading us till very recent times can be framed by them. So, very naturally these hold immense value to people.

Such manuscripts are links in the advancement of people of an area and of humanity, whether for the purpose of technology or for the transmission of information. And though the original had deteriorated, most of these manuscripts were physically copied by scribes, and in effect, the copied manuscript was recopied before reaching us. If people back then were trying to put in such huge efforts to keep the texts alive and maintain their longevity, their hardwork should definitely be not wasted and also preserve them for our own goods and benefits. But being under the weather for years, these documents serve no use unless treated properly. Prevention and digital archiving of historical documents are increasing [23]. These documents suffer from many forms of distortion, degradation, color loss, physical degradation, insect bites, dust, disasters, low-quality paper, etc. Archived images required enhancement and restoration [14] using various techniques to obtain its original look. Restoration helps to extract text and many features from images and helps us preserve manuscripts. Fig.1 depicts the sample of a damaged Indian Manuscript which can be corrected with our model.



Fig 1: A Degraded Manuscript with irregular holes.

Image inpainting is a process of filling up degradation in images or reconstructing them. Earlier it was done by humans but now Deep Learning techniques have replaced them. It produces images such that the existence of damage wouldn't be understood. In these methods, the input is required as a mask showing the regions that should be inpainted. So, old manuscripts require filling up irregular holes in them, in order to form documents from which historians can infer useful data. Not a lot of Deep Learning techniques exist which tackle the problem of image restoration. One approach is using stacked up partial convolution functions while the mask is being updated [1]. Restoration of old manuscripts through inpainting methods will result in a more accurate restoration of manuscripts.

This paper tries to achieve the segmentation of irregular holes using the U-Net [8] approach. Olaf Ronneberger [8] used U Net for the segmentation of neuronal structure during the ISBI challenge. He shows that the network can be trained from a few images by strong use of data augmentation. The architecture is fast and outperforms which shows precise localization. It can also be defined as a neural network designed as an encoder-decoder which solves end to end segmentation tasks. There exists an algorithm related to supervise U-Net. This method is utilized for automatic segmentation in the field of medicine [3].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:14:04 UTC from IEEE Xplore. Restrictions apply.

Apparently UNet was optimized to have importance in the life sciences genre. UNet also finds use in Iris Segmentation. Iris Segmentation has its need in medical sectors as well as in the biometric division. The plus point of using UNet here is its ability to comprehend accurate models from a rather small dataset [4].

Apart from UNet, Partial Convolution [1] covers a significant part of this paper which achieved an impressive result in image inpainting. Dynamic Pacemaker Artifact Reduction [5]. Another problem solved by Partial Convolution (in place of traditional convolution methods) is the restoration of digital Dunhuang murals. It forms a partial convolution layer with the help of the partial convolution operation and masks update function [6]. Image to image convolution neural networks [24] form the basis of Metal Artifact Reduction (MAR) techniques. This problem is seen in painting as a direct relation to a sonogram correction issue and hence partial convolution with UNet architecture is allowed to be used in a part of the pipeline [7].

The major task includes collecting the damaged dataset from all the possible resources and having wide variations, training the model using deep learning, and then testing it. Combining the UNet and the Partial Convulsion architectures, a new solution to the restoration of irregular holes in degraded documents was devised. This method is seemingly new and it has not been explored before this. A mask of the irregular hole will be primarily created in the degraded document and then inpaint them using Partial convolution.

The organization of this paper is as follows. In the section that follows a literature review is provided. In the next section, detailed information related to the dataset is presented to deeply understand the problems that will be dealt to correctly identify the damages done and inprint the manuscript. A detailed description of the architecture employed to design the algorithm is then presented. The process shall continue until the most accurate results possible are obtained for the model. The results and conclusion are then mentioned at the end of the paper.

II. LITERATURE REVIEW

Over the years studies and experimentation methods to extract text from a variety of handwritten, degraded, or historical documents have surfaced. Recently it has been suggested that low contrast texts from the degraded documents can be figured out from the intersection operation of Laplacian and Sobel edge images. [15] All pixels around the edges are enhanced by the Laplacian Operation. Noise pixels are produced, as a result of background variation. On the other hand, the Sobel operator carries out the enhancement of the high contrast pixel without generating noise pixels. The intersection of the operations, result in sieving out the pixels which are relevant for text line segmentation. Further, the width of the edge pixels is reduced by the use of the skeleton. A clustering based on the nearest neighbor criterion is used to segment the text lines. The cluster with lesser branches is the text cluster. This method was devised for transcription of documents from the Indus Valley Civilization. Document image binarization is a useful approach, to separate the background from the foreground by attaching 0 and 1 values to ink and non-ink elements. The most recently developed method to achieve the same is BiNet [16]. This method has been inspired by the U-Net architecture. It uses deep-encoder decoder networks. Here the function is a mapping from the input image to the output image:$f(x) = x + \Delta$, where x is the DSS (Dead Sea Scrolls image collection), image input, and $\Delta$is the noise (information other than the original content). The classification error of the loss function is then minimized and a regression line is plotted for the function to give the prediction.

Another one is the OCR less approach [17]. It proceeds with a variety of image processing techniques followed by de-skewing of images. Image Quality metrics are used for the evaluation of the obtained original text document from the degraded input. The process of deskewing detects and extracts the SURF feature. The geometric transformation between skewed and deskewed images is estimated. The image is converted to Grayscale from the RGB, using the format YCbCr for image restoring and then normalization is applied. In this technique, luminance comparison is done and the values are copied from the RGB image to the Grayscale image.

Restoration is another important aspect of the digitization of degraded documents. Initially global and local thresholding methods are applied. The Niblack method is used to calculate the local threshold of each pixel. $T(x,y) = m(x,y) + k * s(x,y)$, where m is the average of the local area, and s is the standard deviation, and k adjusts the boundary. Acute changes to the spatial averaging are done by the Gaussian Filtering Method [18]. The sigma of the Gaussian distribution gives the degree of noise reduction. Gabor transformation accounts for the preprocessing task, followed by the preprocessing done by the Gaussian Filtering Method. Finally, the edges are detected and the holes are filled. The grey level gradient at a pixel is estimated to make the background uniform and remove the stains.

Another approach makes use of the Markovian-Bayesian clustering [19]. The process is composed of three steps. The image is modeled by 4-nearest neighbors. In the first step, parameters of the Gaussian model are estimated by the iterative conditional estimation (ICE). Hence the result proves the over-segmentation behavior, by dividing the region into numerous regions. The next step in the algorithm is the regions merging based on the degree of homogeneity. Bhattacharya distance is used in this regard. Finally, binarization with k-means clustering is done which goes on to achieve text segmentation of the document.

Multispectral imaging [20] has also gained some prominence in the field of document restoration. This method is based on the concept of pattern recognition. The proposed model is an unsupervised restoration model. The iron-gall ink is chosen as it gives different results in infrared channels and the degradation continues to be visible. When the degradation and the text overlap the end-members extraction is used to

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:14:04 UTC from IEEE Xplore. Restrictions apply.

isolate the text pixels. After removing these pixels from the mask, the inpainting is then applied to the refined mask thus formed. This method effectively preserves the quality of the document as well as the original view.

Decompose algorithm is another approach of looking into the same problem, where it is started by detecting the presence of a Bimodal Histogram [21]. If the answer is affirmative the global thresholding method (time complexity being O(n)) with the highest accuracy is applied to it. This method is vividly effective in case there isn't any non-bimodal histogram present. It then focuses on the 8-bit grayscale of the document. Local feature vectors are then used for thresholding a small area. These local regions are then classified and their edges are smoothened. Finally, thresholding is applied to each region. In the case of the absence of bimodality in the image, a recursive decomposition approach takes place using quadtrees. A thresholding method based on the weighted gradient is used once the entire image has been decomposed. The complexity of this one is $O(n^2)$.

Another approach is to divide or segregate a degraded image into the background, original text, and interfering text (which is then replaced by the average of the background pixel) [21]. A segmentation approach is applied recursively on decorrelated data as it is most unlikely to achieve the result by single clustering. On generated data, recursion is applied to Principal Component Analysis (PCA) and K-means algorithms. PCA is a dimensionality reduction algorithm based on eigenvectors. The k-means algorithm is used which divides the image pixels into two classes.

Mostly, the restoration problem is an unsupervised learning problem and newer algorithms are coming up. But to date, there isn't any structured algorithm to completely solve the degraded document issue. This has been primarily aimed in our paper.

III. METHODOLOGY

The aim is to inpaint the holes or damaged parts inside a document. A two-stage process is presented. The first stage does segmentation of irregular holes and creates a mask using U Net. The mask is then dilated [14]. To remove the boundary of irregular holes. The Second stage does Partial convolution which takes an image and a mask as input and produces a new reconstructed Image. The fig.2 depicts the architecture of our model.

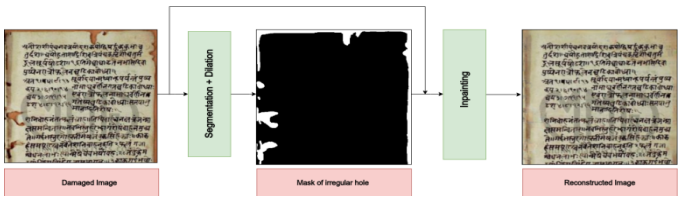


Fig.2: Architecture of our model

3.1 *Dataset*

The role of the dataset in this paper is massive. The primary source of the dataset was the internet where a few sites provide an open and usable database of manuscripts. Apart from the Internet, the ancient Indian Manuscripts are collected from museums, history textbooks, and from the Library of Delhi Technological University. It contains Sanskrit, Prakrit, Gurumukhi, Avadhi Languages along with the Medieval Paleographic Scale (MPS) [9] dataset.

The quality of the documents varied ranging from different handwritings to different page qualities all of which were taken into account during the preprocessing of the images of the manuscripts. The images are resized and cropped into 512 x 512. The dataset is categorized into the damaged and non-damaged dataset. The damaged dataset is used to create a mask and then train the UNet Network. The non-damaged dataset and random mask generator are used to train the partial convolution network. Data Augmentation technique is applied to the dataset to generate more document images, to help train the network with fewer images. Table-1 shows the data in tabular form.Fig.3 shows the sample images of the dataset.

Table 1: Description of dataset

Language	No.of Images	Non-Damaged	Damaged
Sanskrit	1200	896	304
Prakrit	1200	923	277
Punjabi	500	402	98
Avadhi	250	205	45
MPS dataset	2500	1986	514



a) Prakrit b) Sanskrit c) Avadhi d) Punjabi e) MPS

Fig.3: Sample images from the Non-Damaged dataset

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:14:04 UTC from IEEE Xplore. Restrictions apply.

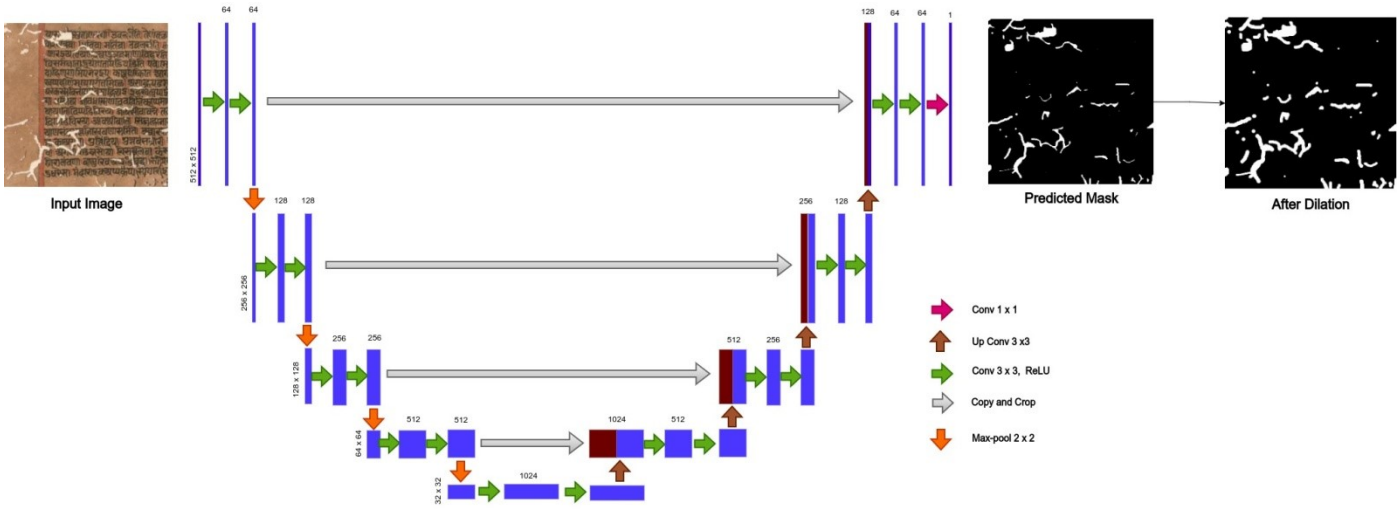


Fig.4: UNet Architecture of our model

3.2 *UNet*

The U-Net architecture [8] is a U-shaped Fully Convolutional Network (FCN) that results in the precise segmentation of the image. The improvement on FCN with the emergence of UNet is that UNet is a symmetric design and there is a skip connection between the downsampling paths. Moreover, the operator used in the upsampling path is concatenation and not sum. It is a symmetric network that is made up of downsampling and upsampling paths. It also contains a great number of feature maps to pass the required information via the upsampling route. UNet has its own advantage of not having dense layers. It makes it suitable for various sizes of image input. Massive data augmentation is also processed efficiently in the UNet.Fig.4 depicts the UNet Architecture. UNet architecture consists of three parts:

A. The contracting/downsampling path:

Recursive application of convolution creates the contracting path. This path contains 4 blocks. Each block has two 3x3 Convolution Layers + ReLU activation functions (with batch normalization), and a 2x2 Max Pooling layer. After the execution of every pooling functioning process, the value of the number of feature maps increases by double. The pooling layer ensures that the number of channels going to the input matrix is always fixed. There is an increase in feature information and a decrease in the spatial ones while contraction takes place. The intention here is to be familiar with the context of the image that is given as input. This enables a better segmentation. This information is then passed to an upsampling way.

B. Bottleneck:

The intention here is to be familiar with the context of the image that is given as input. This enables a better segmentation. This information is then passed to an upsampling way.

C. The expanding/upsampling path:

A combination of the feature and spatial information occurs in the expansive path. It happens by sequential execution of up-convolutions and concatenation operations. The upsampling part also contains 4 blocks. Each block has a Deconvolution layer with two 3x3 Convolution layers + ReLU activation function (with batch normalization).

As a result of this, the upsampling path is almost symmetric to the downsampling path and gives us a u-shaped architectural structure. The task of the U Net architecture is to perform the classification of each pixel of the input image and label it. The final output consists of pixel-wise softmax-max of the entire final feature map with a cross-entropy loss function.

The soft-max is expressed as:

$$p_k(x) = \frac{exp(a_k(x))}{\sum_{K'=1}^{K} \exp(a_{k'}(x))} \quad (1)$$

$a_k(x) = activation\ in\ feature\ channel\ k\ at\ the\ pixel\ position$
$K = Number\ of\ classes (in\ our\ model\ K = 2)$
$p_x(x) = approximated\ maximum\ function$

The cross-entropy at each position $p_{i(x)}(x)$ is written as:

$$E = \sum_{x \propto \Omega} w(x) \log(p_{l(x)}(x)) \quad (2)$$

$\Omega \rightarrow \{0,1\}\ true\ label\ of\ each\ pixel$

Olaf Ronneberger [8] highly uses data augmentation with few mages that help to train the model with few datasets. There is also a few irregular images. So, this research work has used the UNet network and highly used data augmentation to train our model.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:14:04 UTC from IEEE Xplore. Restrictions apply.



Fig. 5: i)The first row shows damaged images ii) The second row shows the mask of irregular holes iii) The last row shows the reconstructed image.

3.3 Partial Convolution

This model consists of stacked partial convolution operations and updating masks to achieve image inpainting. A partial convolution along with a mask update in place of traditional convolution layers gives a much better result in image inpainting involving irregular holes. The partial convolution operation is expressed as:

$$x' = \begin{cases} W^T(X \odot M)\frac{1}{sum(M)} + b\,, & if\ sum(M > 0) \\ 0, & otherwise \end{cases} \quad (3)$$

$W = convolution\ filter\ weights\ for\ the\ convolution\ filter$
$b = corresponding\ bias.$
$X = the\ feature\ values\ (pixels\ values)\ for\ the\ current\ convolution\ (sliding)\ window.$
$M = is\ the\ corresponding\ binary\ mask.$
$\odot = element - wise\ multiplication.$

The output values depend only on the unmasked inputs. The mask is updated after each partial convolution operation. The mask M is updated as:

$$m' = \begin{cases} 1, & if\ sum\,(M) > 0 \\ 0, & otherwise \end{cases} \quad (4)$$

The architecture is similar to U Net, restoring all convolutional layers by the partial convolutional layer and nearest neighbor up-sample during decoding stages. The proposed work has used the same architecture as proposed by Guilin Liu [1]. According to [1] loss function is express as:

$$L_{total} = L_{valid} + 6L_{hole} + 0.05L_{perceptual} + 120(L_{style_{out}} + L_{style_{comp}} + 0.1L_{tv}) \quad (5)$$

$L_{total} = Total\ loss$
$L_{valid}, L_{hole} = Per\ pixel\ loss\ of\ the\ image$
$L_{style_{out}}, L_{style_{comp}} = Style\ loss$
$L_{tv} = Total\ variantion\ loss$

IV. EXPERIMENTS AND RESULTS

The dataset is further divided into three-parts-training, validation, and testing. 60 % of data was used for training, 20 % for validation, and 20 % for testing. All the images are of size 512 x 512.

4.1 Masking

As shown in fig 1 and fig 5 the color of irregular holes is mostly white. The colored damaged image is converted into grayscale. The Non-hole part pixels in an image are low and the holes in the images have high pixels. So a threshold point is randomly selected and this helps in the binarization of the image. The different images have different threshold points. Some holes are not white in color but their pixel value is higher than the non-damaged parts of the image.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:14:04 UTC from IEEE Xplore. Restrictions apply.

4.2 Training Process

This process has used damaged images and their manually annotated masks to train the U Net model using Adam and binary cross-entropy with up to 200 epochs. For training the Partial Convolution, random masks of size 512 x 512 are generated using OpenCV consisting of rectangles, circles, lines. Further this process will use these masks and the non-damaged dataset to then train the partial convolution network. This research work has started with the initial weight [11] and Adam [12] Optimization. Holes create a problem for Batch Normalization due to the mean and variance computed for hole pixels. To solve the mean and variance problem, the model is trained in two steps, which also helps to gain faster convergence. The model is trained in two steps first with initial training with enabled Batch Normalization with the learning rate of 0.0002 with up to 120 epochs and then fine-tuning with disabled Batch Normalization in the encoder with the learning rate of 0.00005 with up to 80 epochs. The batch size is 5 and 2500 steps per epoch in both cases. Fig.6 shows the corresponding results with the predicted output after masking on non-damaged dataset images with random lines, circles, and rectangles.



Fig.6: i) Masked Image ii) predicted image iii) Original image

4.3 Performance on the degraded dataset

After Training, the proposed model is tested on the damaged dataset to inpaint its irregular holes. The proposed model is tested for every angle of damaged images and for small irregular holes to larger irregular holes. Fig.5 shows the damaged images, their masks, and the corresponding predicted outputs.

4.4 Qualitative Analysis

Different square masks of sizes-16 x 16, 32 x 32, 64 x 64, 128 x 128 and straight-line masks of pixel 4 x 512, 8 x 512, 16 x 512, 32 x 512 on a 512 x 512 sized image are created. These masks are then added on non-damaged images and then image prediction is done in our model. Two Quantitative approaches are used to compare the predicted image and original image using Peak Signal to Noise Ratio(PSNR) which shows the ratio of the actual signal and the noise in the image, and Structural Symmetry Index(SSIM) which tells the similarity index of two images as shown in Table 2 and 3.

Table 2: PSNR AND SSIM values for each of the square masks

	16 x 16	32 x 32	64 x 64	128 x 128
PSNR	38.17	37.46	32.528	26.05
SSIM	0.972	0.971	0.965	0.935

Table 3: PSNR AND SSIM values for each of the straight-line masks

	4 x 512	8 x 512	16 x 512	32 x 512
PSNR	36.83	35.85	34.67	29.78
SSIM	0.971	0.966	0.964	0.94



Fig.7: i) Damaged image ii) Damaged mask iii) Reconstructed image

V. DISCUSSION

This paper proposes the U Net and partial convolution model to create an automatic mask and then automatic mask update to inpaint it. This model can easily handle small irregular holes, small lines, and recover the text

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:14:04 UTC from IEEE Xplore. Restrictions apply.

with utmost perfection. But, it somehow does not show the best results and fails to recover text and inpaint the larger holes. Fig 7 shows some error due to dilation and not proper segmentation of irregularly large holes.

VI. CONCLUSION

This paper is a novel initiative to help restore degraded old Indian Manuscripts using UNet and partial convolution techniques. The specific work is done in this paper ie. of inpainting holes in a manuscript using Deep learning algorithms show promise with regard to computational performance. Concerning the accuracy, the network did perform well. It was efficiently able to inpaint irregular holes and tears of small sizes. However, it did face some difficulty in tackling larger holes and damages due to the dilation of the images. On using different datasets the architecture applied still proves to be very accurate every time.

VII. FUTURE WORK

IN future, it will include a wider dataset and irregular mask dataset [1] for our model to make it more robust. Also, the problem of error due to dilation that occurred in manuscripts along with larger holes will be catered to, in order to make the digitization of manuscripts easier and efficient. Nevertheless, it will also use other inpainting and segmentation algorithms to make the model to remain more accurate.

Acknowledgement

We would like to thank Delhi Technological University and our friends and family who helped us in finalizing the paper. We are sincerely thankful to www.indianmanuscripts.com from where we have collected most of our dataset.

References

[1] Guilin Liu, Fitsum A. Reda, Kevin J. Shih, Ting-Chun Wang, Andrew Tao, Bryan Catanzaro, Image Inpainting for Irregular Holes Using Partial Convolutions. The European Conference on Computer Vision (ECCV), 2018, pp. 85-100.

[2] Falk, T., Mai, D., Bensch, R., Çiçek, Ö., Abdulkadir, A., Marrakchi, Y., Böhm, A., Deubner, J., Jäckel, Z., Seiwald, K., Dovzhenko, A., Tietz, O., Bosco, C.D., Walsh, S., Saltukoglu, D., Tay, T.L., Prinz, M., Palme, K., Simons, M., Diester, I., Brox, T., Ronneberger, O., 2018. U-Net: deep learning for cell counting, detection, and morphometry.

[3] Li, S., Tso, G.K., He, K., 2020. Bottleneck features supervised U-Net for pixel-wise liver and tumor segmentation. Expert Systems with Applications 145, 113131.

[4] Lozej, J., Meden, B., Struc, V., Peer, P., 2018. End-to-End Iris Segmentation Using U-Net. 2018 IEEE International Work-Conference on Bioinspired Intelligence (IWOBI).

[5] Elss T.L., N., Nickisch, H., Wissel, T., Morlock, M., Grass, M., 2020. Learning metal artifact reduction in cardiac CT images with moving pacemakers. Medical Image Analysis 61, 101655.

[6] Chen, M., Zhao, X., Xu, D., 2019. Image Inpainting for Digital Dunhuang Murals Using Partial Convolutions and Sliding Window Method. Journal of Physics: Conference Series 1302, 032040.

[7] Samoylenko, A., Golanov, A., Pimkin, A., Antipina, N., Ovechkina, A., Daechina, A., Belyaev, M., 2019. Multi-Domain CT Metal artifacts reduction using partial convolution-based inpainting.

[8] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In MICCAI, pages 234–241. Springer, 2015

[9] Sheng He, Petros Samara, Jan Burgers, Lambert Schomaker. Multiple-Label Guided Clustering Algorithm for Historical Document Dating and Localization IEEE Trans. on Image Processing, Vol. 25(11), Nov. 2016. http://ieeexplore.ieee.org/document/7551181/

[10] Jonathan Long, Evan Shelhamer, Trevor Darrell, UC Berkeley, Fully Convolutional Networks for Semantic Segmentation.

[11] He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE international conference on computer vision. pp. 1026–1034 (2015)

[12] Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)

[13] M.Raid, al, Image Restoration Based on Morphological Operations, International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), Vol. 4, No.3, June 2014.

[14] Yahya, S. R., Abdullah, S. N. H. S., Omar, K., Zakaria, M. S., Liong, C. Y., "Review on image enhancement methods of the old manuscript with damaged background," in International Conference on Electrical Engineering and Informatics, pp. 62-67, 2009.

[15] Kavitha, A., Shivakumara, P., Kumar, G., Lu, T., 2016. Text segmentation in degraded historical document images. Egyptian Informatics Journal 17, 189–197.

[16] Dhali, M.A., Wit, J.W., Schomaker, L. Binet: Degraded-Manuscript Binarization in Diverse Document Textures and Layouts using Deep Encoder-Decoder Networks. [eprint] 2019.

[17] Hamsaveni L., Prakash N., Suresha., 2017. Degraded Document Analysis and Extraction of Original Text Document: An Approach without Optical Character Recognition. International Journal of Computer and Information Engineering 11, 110-114.

[18] Raj V., Arunkumar C., 2014. Content Restoration of Degraded Termite Bitten Document Images. International Journal of Computer and Information Engineering 4, 151-155.

[19] Hedjam R., Moghaddam R. F., Cheriet M., Text Extraction from degraded document images, 2010 2nd European Workshop on Visual Information Processing (EUVIP), Paris, 2010, 247-252.

[20] Hedjam, R., Cheriet, M., 2013. Historical document image restoration using a multispectral imaging system. Pattern Recognition 46, 2297–2312.

[21] Chen, Y., Leedham, G., 2005. Decompose algorithm for thresholding degraded historical document images. IEEE Proceedings - Vision, Image, and Signal Processing 152, 702.

[22] Drira, F., a. n.d. Towards Restoring Historic Documents Degraded Over Time. Second International Conference on Document Image Analysis for Libraries (DIAL06).

[23] Apostolos Antonacopoulos, Andy Downton, Special issue on the analysis of historical documents, International Journal on Document Analysis and Recognition 9 (April (2)) (2007) 75–77.

[24] Vijayakumar, T. (2019). Comparative study of the capsule neural network in various applications. Journal of Artificial Intelligence, 1(01), 19-27.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:14:04 UTC from IEEE Xplore. Restrictions apply.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/343851471

Introduction to Microwave Remote Sensing

Preprint · August 2020

DOI: 10.13140/RG.2.2.31159.55205

CITATIONS
0

READS
18

1 author:



Sahaj Rastogi
Delhi Technological University
1 PUBLICATION **0** CITATIONS

SEE PROFILE

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

Intoduction to Microwave Remote Sensing View project

All content following this page was uploaded by Sahaj Rastogi on 25 August 2020.

The user has requested enhancement of the downloaded file.

Introduction to Microwave Remote Sensing

Sahaj Rastogi, sahajr16@gmail.com

***Abstract*—This paper reviews a brief overview of the Microwave Remote Sensing and the History of Microwave Remote Sensing along with its instruments such as Radiometers (Passive Sensors) and Imaging Radar (Active Sensors). This paper is inspired by various other studies based on the history and launches of various satellites for remote sensing using radiometers and imaging radar with their pros and cons.**

***Index Terms*— CORONA, IKONOS, Imaging Radar, IRS-1A, Microwave, Radiometer, Remote Sensing, SEASAT**

I. Introduction

THE word Remote Sensing is defined as the process of measuring the physical properties of any distant objects typically from aircraft and satellites using reflected or emitted energy.

Remote sensing collects data about Earth by detecting the energy that is reflected from the surface. These sensors are mounted on either UAV/drones which are close range or Aerial or Satellite sensors which are far range. Microwaves are a kind of electromagnetic wave whose wavelengths starting from about 1 cm to 1 m; with frequencies between 3 GHz to 30 GHz. The microwave ranges are often referred to as the different frequencies bands: S, C, X, Ku, K, Ka (IEEE radar band). Photography had set the beginning of the history of remote sensing. The origin of other types of remote sensing can be traced back to World War II, with the development of radar, sonar, and thermal infrared detection systems.

TABLE 1
VARIOUS IEEE MICROWAVE RADAR BANDS AND THEIR FREQUENCY AND WAVELENGTH RANGES.

Microwave Band	Frequency	Wavelength
S band	2 to 4 GHz	7.5 cm to 15 cm
C band	4 to 8 GHz	3.75 cm to 7.5 cm
X band	8 to 12 GHz	25 mm to 37.5 cm
K_u band	12 to 18 GHz	16.7 mm to 25 mm
K band	18 to 26.5 GHz	11.3 mm to 16.7 mm
K_a band	26.5 to 40 GHz	5.0 mm to 11.3 mm

29th July 2020. This work was supported with the help of Delhi Technological University, India under MCG Internship 2020.

Microwaves were chosen for remote sensing applications due to their numerous advantages. This part of the spectrum has unique capabilities for remote sensing objects in the day as well as night. These waves can penetrate through clouds and penetrate, acquiring images in any weather. Not only they can transmit through water but can penetrate vegetation canopy which is useful for studying surface water or flood extent.

II. History of Remote Sensing

Remote sensing is generally referred to as the use of aircraft or satellite-based technologies to collect information to detect and classify objects on Earth.

The history of remote sensing started with the famous balloonist G. Tournachon, who made photographs of Paris from his own balloon in 1858. Earlier ways of communication were pigeon messenger and other instruments like kites, rockets. The need for systematic aerial photography was required during World War 1 for military surveillance. Various modified aircraft like P-51, OV-1, and many others were used to collect information for aerial photography. During the war, the cameras were specially designed for aircraft use and were mounted on the aircraft. The data obtained from the aircraft sensor pods were used by the government and the military which later was also accessible for the public in the late 20th century.

Earlier the term "aerial photography" was used instead of "remote sensing" which was initially introduced in 1960 due to the advancements in aerial photography beyond the traditional black and white photos. In the period of 1960-1970, there was a drastic shift to carry out remote sensing from satellites rather than airplanes, as satellites tend to cover a greater area than planes and ensure regular monitoring. A major change from analog to digital storing of imagery was occurring, which made it possible to analyze and display images on computers. Computer technology in the same era was advancing from large mainframe computers to compact microcomputers which enabled graphic form rather than numerical outputs. Later in the second half after the Cold War, there was a visible development of artificial satellites which allowed remote sensing to become globally accessible. These sensors due to advancements were now able to record several different portions of the electromagnetic spectrum, which is beyond what the human eye could view. Several things occurring on the Earth's surface were now visible which were previously inaccessible with normal aerial photography.

Until 1995 **CORONA**, America's first satellite series

program was unknown to the people which was declassified by the then-President Bill Clinton. This not only made is accessible to all but also helped scientists to study environmental changes. CORONA was a secretive military reconnaissance program that still continues to serve its mission. Later many different agencies such as National Aeronautics and Space Administration (NASA), Indian Space Research Organization (ISRO), Roscosmos- Russian Space Agency, European Space Agency (ESA), and Japanese Space Agency (NASDA) launched various active and passive sensors in space.

A. SEASAT – 1st Earth-orbiting satellite designed for remote sensing of the oceans.

SEASAT was one of the earliest Earth-observing satellites, launched by Jet Propulsion Laboratory (JPL), NASA with the aim of having a better understanding of Earth's oceans and seas by various oceanographic sensors attached to the satellite. Launched from Vandenberg Air Force Base, California on June 26, 1978.

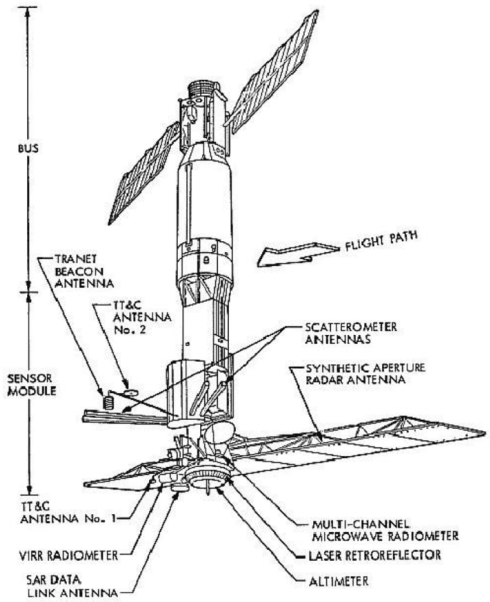


Figure 1. - SEASAT
Parts of the satellite showing its sensors modules, radiometer, altimeter, and different antennas. **(**Image credits © alchetron.com)

The satellite had various instruments which comprised of (i) Radar Altimeter; (ii) Synthetic Aperture Radar (SAR); (iii) SEASAT-A Satellite Scatterometer (SASS); (iv) Visible and Infrared Radiometer (VIRR); and (v) Scanning Multichannel Microwave Radiometer (SMMR).

B. IRS-1A - India's 1st Remote Sensing Satellite System

IRS-1A is the first series of India's indigenous operating remote sensing satellites, which was launched successfully into a polar sun-synchronous orbit on St Patrick's Day, 1988, with the assistance of Soviets from the Cosmodrome at Baikonur.

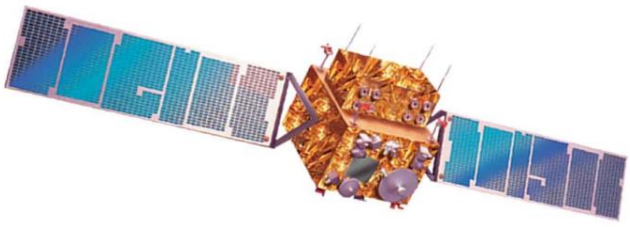


Figure 2. - IRS-1A Satellite
India's first remote sensing satellite (Image credits © Department of Space, ISRO)

The successful launch of IRS-1A not only placed India into the few nations to have their own remote sensing satellite but also made the people of India proud. This showed the maturity of the Indian satellite program advancing with the capability to map the natural resources of its nation.

After the successful launch of IRS-1A, a major milestone in the IRS program, an identical satellite IRS-1B was launched in 1991 which improved the frequency to 11-day cycle. From then onwards, many series of IRS missions with theme-based satellites were launched, namely, (i) RESOURCESAT and RISAT series for Land/Water applications; (ii) OCEANSAT and INSAT for Ocean/Atmospheric studies; and (iii) CARTOSAT for Large scale mapping. These paved ways to improve natural resources management and infrastructure development using microwave remote sensing in the country.

The array of Indian Earth Observation (IEO) satellites that are capable of imaging in visible, infrared, thermal, and microwave regions of the spectrum have provided help in various applications with spatial resolution ranging from 1 km to less than 1 m. spatial resolution ranging from 1 km to less than 1 m.

C. IKONOS - 1st satellite to collect publicly available high-resolution data

IKONOS (Greek word for "Image") was the first commercially available satellite that provided high-resolution satellite sensor data. Launched from the Vandenberg Air Force Base, California, USA on September 24, 1999.



Figure 3. - IKONOS Satellite Sensor
1st satellite to collect publicly available high-resolution data (Image Credits © DigitalGlobe)

IKONOS satellite sensor is operated by DigitalGlobe which provides high-resolution images. It has the ability to capture 3.2 m multispectral, near-infrared 0.82 m panchromatic resolution at nadir. The data collected by IKONOS has applications with include mapping of natural resources, natural disasters, tax mapping, engineering construction and forestry analysis. Generating Digital Elevation Models (DEMs) and 3D Digital Terrain Models (DTMs) to aid in costal monitoring and internal security of nation is also possible due to the high-resolution data available.

III. TYPES OF MICROWAVE REMOTE SENSING

There are two types of microwave sensors used for microwave remote sensing, namely, (i) passive sensors and (ii) active sensors. Passive sensors include Radiometers whereas active sensors include Imaging Radar, Altimeter, Real Aperture Radar (RAR), and Synthetic Aperture Radar (SAR).

A. Radiometers - Passive Sensor

The basic principle of a radiometer is to measure the power picked up by the antenna due to the difference in temperature. Microwave Radiometer measures the naturally emitted radiation, the brightness temperature of substances which are within its bandwidth, and antenna's field of view. The radiometer is a calibrated and sensitive microwave receiver.

Since all bodies radiate power above absolute zero (0 K = -273°C), according to Plank's Law. While considering microwave frequencies the Rayleigh-Jeans holds the approximation that radiation is proportional to the physical temperature of the body. This is for the ideal case that is black bodies but for natural bodies that are not perfect emitters, the term emissivity is introduced, from 0 to 1, which tends to describe how well the body radiates with respect to the blackbody. The brightness temperature of a blackbody is thus adequate to its physical temperature; while-all natural bodies will have brightness temperatures less than that.

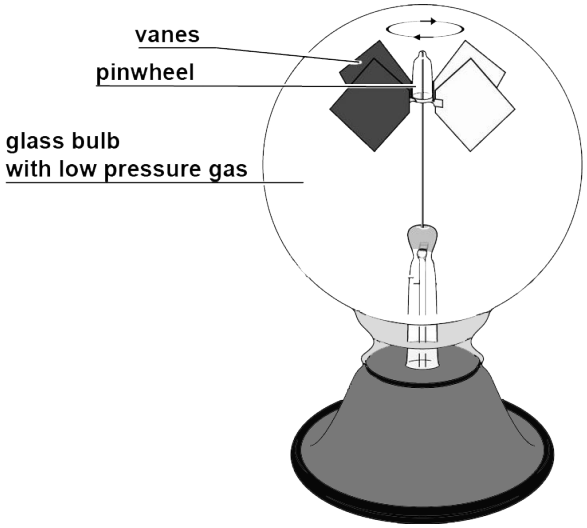


Figure 4. - Radiometer
Upon flashing a light source the black side starts to absorb more energy than the white side counterpart, which creates a temperature difference between both sides, so the pinwheel starts to move. (Image credits © physicsexperiment.co.uk)

For example, ice is close to a blackbody at microwave, thus it will have T_d close to (=260 K) while water has T_d in the range of (=150 K). The contrast in the two different temperatures is why spaceborne radiometry is being used to monitor sea ice with such great success. Various types of Radiometers include (i) Total Power Radiometer – Simplest of all with optimum sensitivity used in medial sensor applications; (ii) Dickle Radiometer – Improved version of Total Power Radiometer to increase stability used in instruments for detecting and measuring radiant energy; and (iii) Correlation Radiometer – Most advanced which consists of two Total Power Radiometer where the signals before detection are cross related to producing two additional outputs used in remote sensing.

B. Imaging Radar - Active Sensor

An Imaging Radar can be imagined as a camera with a flashlight that provides its own illumination to light up an area but at radio wavelengths. Instead of using a camera lens, radar uses an antenna and digital computer to record the images.

For imaging radar, very high power pulses per second are transmitted, with each pulse duration between 10-50 microseconds (μs) towards the target or imaging area.

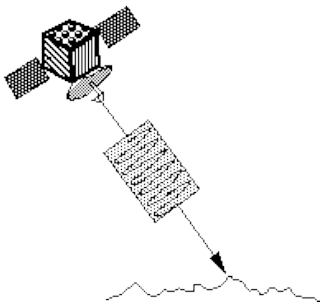


Figure 5. – Imaging Radar transmission
Radar's transmitted pulse is reaching the surface of Earth.

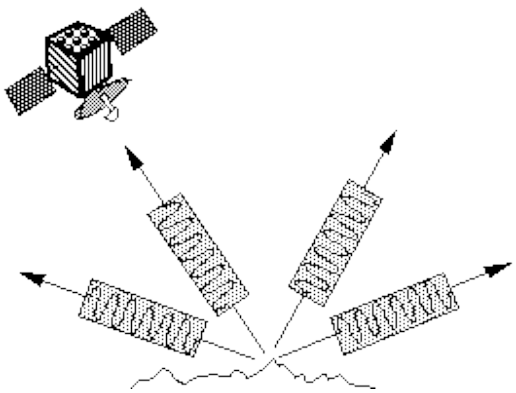


Figure 5. – Imaging Radar backscatter
Satellite measures reflected echo which is also known as backscatter.

In the case of imaging radar, the satellite moves along a flight path with the illumination of the area using the attached radar. Imaging Radar involves a different number of ways to present

radar signals. Starting with; (i) A-scan which is a graph between amplitude vs. time scan, unable to provide information about the direction of layers; (ii) B-scan is obtained by converting the graph of A-scan into brightness and then displayed on the screen; and (iii) P-display or Plan Position Indicator (PPI) used the converted A-scan information in the same relative directions as the antenna orientation. The radar screen shows the surrounding picture from the center of the screen. Plan Position Indicator is the most widely used instrument in navigation of ships and aircrafts.

IV. Conclusion

The study enabled us to learn about the rich history of our past achievements on Microwave Remote Sensing. In the future, we may further use Artificial Intelligence and Machine Learning to advance in Microwave Remote Sensing with greater shifts to observe. Finally, I would say that although Imaging Radar has proved to be the best but are very costly while passive sensors are cheap due to non-usage of own illumination.

Acknowledgment

I would like to thank Delhi Technological University for providing this opportunity to research and publish a review paper on the Introduction and history of microwave remote sensing.

References

[1] P. S. Roy, M. D. Behera, S. K. Srivastav, "Satellite Remote Sensing: Sensors, Applications and Techniques," in Springer, DOI: 10.1007/s40010-017-0428-8, [Online]

[2] GERALD K. MOORE (1979) What is a picture worth? A history of remotesensing / Quelle est la valeur d'une image? Un tour d'horizon de télédétection, HydrologicalSciences Bulletin, 24:4, 477-485, DOI: 10.1080/02626667909491887, [Online]

[3] Calla, O. P. N. (2008). "Microwave remote sensing using space borne sensors". 2008 International Conference on Recent Advances in Microwave Theory and Applications. doi:10.1109/amta.2008.4763255, [Online]

[4] Skou N. (2008)" Microwave Radiometry and Radiometers for Ocean Applications". In: Barale V., Gade M. (eds) Remote Sensing of the European Seas. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-6772-3_16, [Online]

[5] Remote_Sensing, "https://en.wikipedia.org/wiki/Remote_sensing #History"

[6] https://crisp.nus.edu.sg/~research/tutorial/mw.htm

[7] https://oceanservice.noaa.gov/facts/remotesensing.html

[8] https://www.satimagingcorp.com/satellite-sensors/ikonos/



Sahaj Rastogi is currently pursuing B.Tech in Engineering Physics with majors in Electronics from Delhi Technological University, Delhi.
He is passionate about Artificial Intelligence, and Machine Learning.

View publication stats

arXiv:2008.12756v1 [nlin.CD] 28 Aug 2020

Investigating the relation between chaos and the three body problem

T.S.Sachin Venkatesh
tssachin.venkatesh@gmail.com
Delhi Technological University, Delhi 110042, India

Vishak Vikranth
Bengaluru 560076, India

We review the properties of fractals, the Mandelbrot set and how deterministic chaos ties to the picture. A detailed study on three body systems, one of the major applications of chaos theory was undertaken. Systems belonging to different families produced till date were studied and their properties were analysed. We then segregated them into three classes according to their properties. We suggest that such reviews be carried out in regular intervals of time as there are an infinite number of solutions for three body systems and some of them may prove to be useful in various domains apart from hierarchical systems.

Key words - Celestial mechanics, Three-body problem, Gravitational interaction, Chaos, Orbits, Astronomical simulations

I. INTRODUCTION

Exploring the connections between different theories of mathematics leads one through successive topics which diverge very little from one another, but looking at the the path as a whole, the initial point of probing and the final point have very little in common. So a quick introduction of the topics we covered is listed below

A. Fractals

Fractals are infinitely complex patterns that are self-similar across different scales. They are created by repeating a simple process over and over in an ongoing feedback loop. A key characteristic of a fractal is its fractal dimension. Unlike the Euclidean dimension, fractal dimension is generally expressed by a non-integer and is an indicator of the complexity or roughness of a given figure. Some common examples include Sierpiński triangle having a Hausdorff dimension of 1.585, Koch's snowflake with a Hausdorff dimension of 1.262 and the coastline of Britain, whose fractal dimension is 1.21.

B. The Mandelbrot set and the logistic map

One of the most interesting fractals, the Mandelbrot set is the set of complex numbers c for which the function $x_{n+1} = x_n^2 + c$ does not diverge to infinity when iterated from z=0. The path of all such orbits when plotted give us an image which is intuitive and rather easy to understand. All the points inside the main cardioid have a single fixed point, all the points inside the main bulb have 2 limit points and subsequently all other bulbs represent the set of numbers which have different number of limit points. The logistic map is a polynomial mapping of degree 2, often cited as an archetypal example of how complex, chaotic behaviour can arise from very simple non-linear dynamical equations. By simple algebraic manipulation, the logistic map $x_{n+1} = rx_n(1 - x_n)$, can be recoded into the form $x_{n+1} = x_n^2 + c$. This leads us to further explore a relation which produces a rather interesting relation when we change our Point of View. From the branching structure of the logistic bifurcation diagram we can read the cycle number of the corresponding features of the Mandelbrot set.

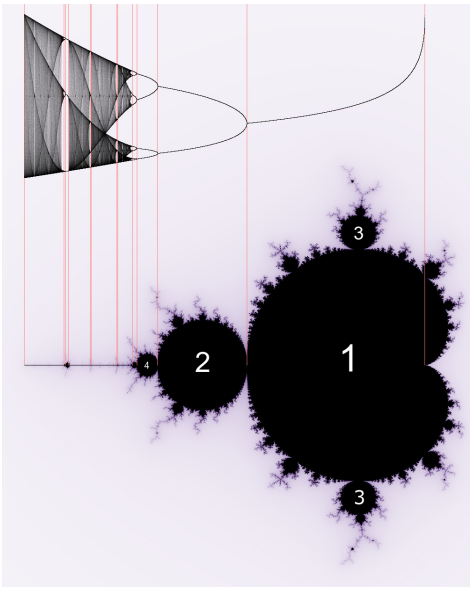


FIG. 1: The real line on the mandelbrot set lines up with the bifurcations in the logistic map[1]

C. Deterministic Chaos

Deterministic chaos is the study of how systems that follow simple, straightforward, deterministic laws can exhibit very complicated and seemingly random long term

behavior. One of the foundations of chaos theory is Edward Lorenz's discovery of systems where the dynamics are sensitive to minute changes in the initial conditions thus are unpredictable. Due to this sensitivity the behaviour of systems appear to be random although model is deterministic, which means that it is well defined and it does not contain any random parameters. In the beginning, this theory was used only in the field of meteorology but soon it was realized that it can also define the other chaotic systems in varied science disciplines which were based on the predictability of future behavior of the systems.

II. THEORY

A. The relation between Logistic Map and Chaos

The logistic equation is of the form $x_{n+1} = rx_n(1 - x_n)$. By changing the values of r, the following behaviour is observed : When r is in $[0, 1]$, the sequence will converge to 0. When r is in $[1, 3)$, the sequence approaches the value $r - \frac{1}{r}$, regardless of the value for x_0;though the rate of convergence of the function decreases as the value of r increases. When r is in $[3, 1 + \sqrt{6})$,the function oscillates between two limit points whose values are determined by r. When r is in $[1 + \sqrt{6}, \sim (3.54409)]$,the sequence will approach permanent oscillations among four limit points. When r increases beyond $\sim (3.54409)$,the function will approach oscillations among eight limit points, then 16, 32,and so on. The ratio between the lengths of two successive bifurcations approaches the Feigenbaum constant $\delta \approx 4.66920$[2]. At $r \approx 3.56995$ is the onset of chaos. We no longer see oscillations of finite period. Slight variations in the initial x_0 value yield drastically different results over time.

For values of r beyond ~ 3.56995 the sequence exhibits chaotic behaviour, but there are still certain values of r that show non-chaotic behavior; these are sometimes called islands of stability. For example for r=$1 + \sqrt{8}$ the sequence oscillates between three values, and for slightly higher values of r oscillation among 6 limit points, then 12 etc.

B. Noise versus Deterministic Chaos

Noise is the random variation of values whereas chaos happens when the initial state variables of the system differ ever so slightly which lead to drastically different outcomes making it impossible to predict the initial value from the output.

We can also differentiate between noise and deterministic chaos using fractals. The Hausdorff dimension of the attractor of a random model is usually infinite. But the attractor of a model of deterministic chaos always gives us a non integral, finite value. Hence, if the Hausdorff dimension of a model is finite it is good indicator that it is a deterministic model that can be represented by a system of non linear differential equations[3].

C. Lorenz Attractor

An attractor is a point, set or even a fractal (strange attractor) around which it the function seems to converge. Edward Lorenz, an American physicist and meteorologist who pioneered chaos theory is famous for demonstrating chaotic motion using "Lorenz water wheel": a water-wheel with holes at the bottom and a constant stream pouring from above. The emptying and refilling of the wheel produces unpredictable changes in its direction of rotation and angular velocity.

Lorenz's paper in 1963,tries to model weather and used these non linear differential equations[4]:

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} &= \sigma(y - x) \\ \frac{\mathrm{d}y}{\mathrm{d}t} &= x(\rho - z) - y \\ \frac{\mathrm{d}z}{\mathrm{d}t} &= xy - \beta z \end{aligned} \tag{1}$$

According to Lorenz's account, while working on the weather model he was running simulations and wanted to repeat a previous simulation. He started the simulation from a random iteration and since it was a computer program it should have given him the exact same results. But the results were in stark disagreement. He initially shrugged this off as an error, but then realized he had not entered the initial conditions exactly. The computer was taking values precise upto six decimal places but the printer only displayed three. He re-entered the rounded off values and this minuscule error caused drastic changes in the outcome.

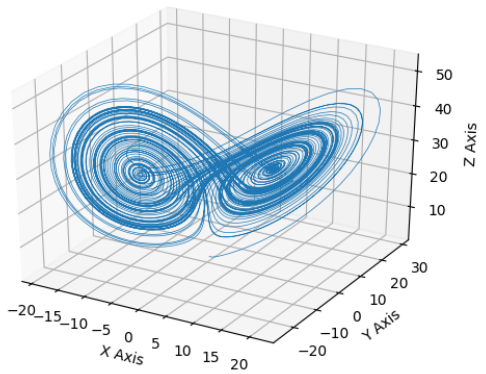


FIG. 2: Lorenz attractor[5]

Lorenz considered the case where $\sigma = 10$, $\beta = 8/3$, $\rho = 28$ with $(x_0, y_0, z_0) = (0, 1, 0)$. This gives us the Lorenz attractor[4].

D. Three Body Problem

The three body problem is a system of ordinary differential equations modelling three bodies of masses under mutual gravitation in two or three dimensions. The motion of such a system was first pondered upon by Newton and was reconsidered by many mathematicians and scientists. The famous three-body problem has had a great influence on physics, mathematics and non-linear dynamics. It has paved way to a new field in modern science, chaotic dynamics.

Until 1975, there were not many models of three body systems and the ones that existed were too restrictive and specific about the parameters of the model. In 1890, Poincaré proved the non existence of the uniform first integral of a three-body problem in general, and also highlighted the sensitive dependence to initial conditions of its trajectories[6]. Three body systems without mass hierarchy are never thought to be stable for very long[7]. They can certainly exist for some period of time, but they aren't found to be long term stable. In these systems, each body orbits the center of mass of the system. Mostly, two of the bodies form a close binary system, and the third body orbits this binary at a distance much larger than that of the orbit of the close binary. This arrangement is called hierarchical. The reason for this behaviour is that if the inner and outer orbits are comparable in size, the system may become dynamically unstable, leading to a body being ejected from the system. And if the system in question consists of disproportionate mass bodies, the Hill sphere[8] mechanism comes into play.

A lot of work has been carried out and various models have been explored such as Poincaré's planar circular restricted three body problem (PCR3BP)[6], Sitnikov models in the early 1900s and the recent discovery of thirteen families of stable planar equal mass three body systems. The recent discoveries have formulated a new method to check for duplicate orbits and sort different instances, they use an abstract space called a 'shape space sphere'[9] which describes the shape of the orbits in terms of the relative distances between the objects. Three spots around the sphere's equator mark where two of the particles would collide, and a line drawn over the ball, which must avoid those spots, maps how near the objects get to each other. One of their original solutions nicknamed 'yarn' is shown.

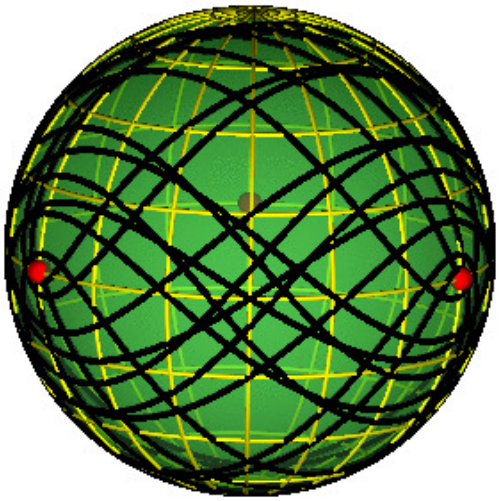


FIG. 3: Shape sphere showing the relative positions of the orbits in a 3-body problem.[10]

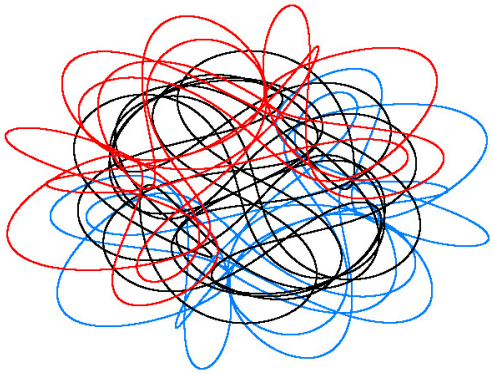


FIG. 4: The same orbits in real euclidean space relative to centre of mass of the system.[10]

III. MODELING AND OBSERVATIONS

We simulated different models of the three body system ranging from the more popular 'Montgomery 8'[11] orbit to custom made orbits like 'perturbed circular orbits'. Most of these models are equal mass systems, but there are some special orbits that were simulated for unequal mass systems, which were observed to remain stable over extensive number of iterations. The simulated models can broadly be classified into three classes of stability; namely, Long-term stable, Quasi-stable and Chaotic systems.

A. Long term stable

The families of such three body systems are extremely rare to compute in a general setup due to the increased complexity in computation owing to the increase in the number of parameters. Some of the first solutions to the three body system were given by Euler and Lagrange for special cases like 'Circular Restricted 3 Body Problem' and 'Planar Restricted 3 Body Problem'. Until 2013, specific solutions could be sorted into just three families: the Lagrange-Euler family, the Broucke-Hénon family, and the figure-eight family. Since then there have been more than a thousand periodic solutions for the planar three body problem and 13 families of solutions to the more

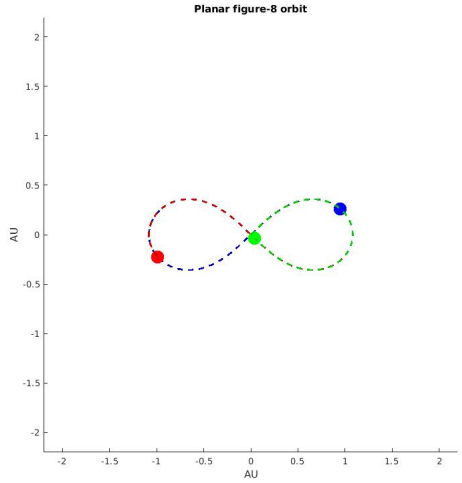


FIG. 5: Montgomery 8 - the system comprises of 3 equal masses with a net angular momentum of zero.

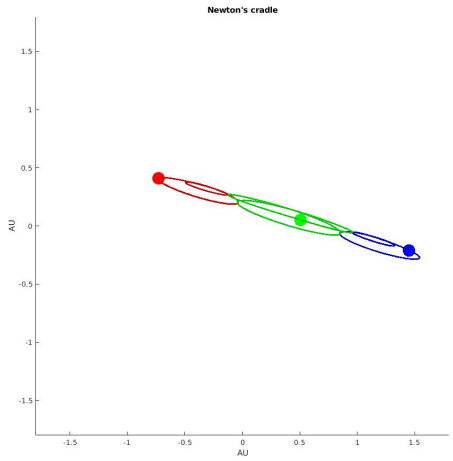


FIG. 6: Newton's cradle-esque system of equal masses that meet at only two points.

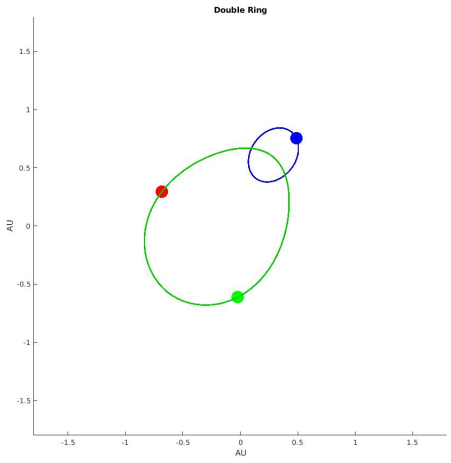


FIG. 7: Double ring system. The blue body acts like a field regulator.

general, non planar three body problem.

The Montgomery 8 orbit uses a variational method to exhibit a simple periodic orbit for the Newtonian problem of three equal masses in the plane[11].

Other stable orbits include Newton's cradle-esque orbit (FIG. 6) which looks as if two binary orbits have coalesced together to form a system with a period of two. There are two separate levels of orbital systems intertwined into one. The bodies try to slingshot each other but owing to their gravitational pull, they aren't able to escape the orbit, thus when they try to slingshot, they form a larger orbit but when they are unable to exit the system, they are dominated by the gravitational potential and form a shorter orbit.

A more rational system that has a higher probability of being detected in the universe is the system shown below (FIG. 7). This is a solution for an equal mass system where two bodies move in the same orbit while the third body acts as the field regulator, making sure the other two bodies stay in the same orbit. The probability of the existence of such system is relatively high due to the simplicity of the system and the lack of complex orbits which are very sensitive to even the slightest of perturbations.

B. Quasi-stable

Three body systems corresponding to this family are much harder to find since these systems remain stable for some period of time which may range anywhere from a

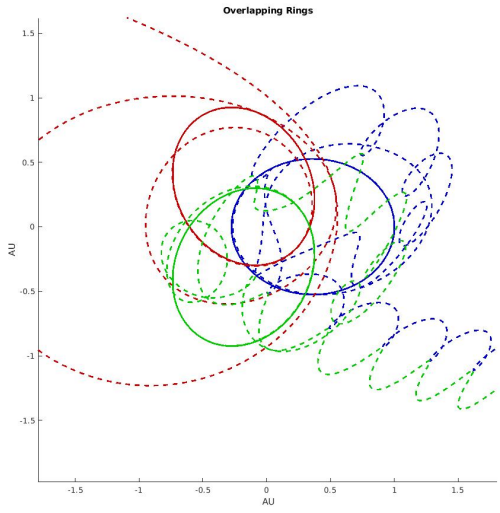


FIG. 8: Overlapping rings with a common centroid. The equal mass bodies are placed at the vertices of an equilateral triangle.

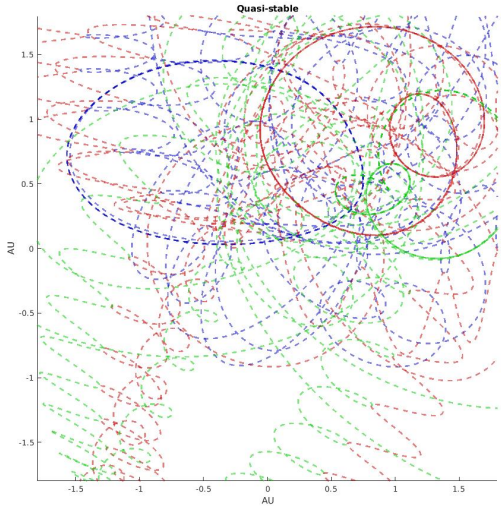


FIG. 9: Unequal mass system destabilised due to perturbations.

FIG. 10: The highlighted curves represent the quasi-stable orbits of the systems, while the dashed lines represent the path of the bodies after destabilizing.

few moments to a few hundreds of years but over longer periods of time, they seem to diverge from their orbits and the three body system breaks down to a two body system with the third body being flung away. Such systems are perfect examples to demonstrate sensitive dependence to initial conditions (SDIC), even the slightest changes in the initial conditions of positions, velocity, mass can alter their course and send them travelling in a completely different direction or configuration.

One such system is where the equal mass bodies are initially placed at the vertices of an equilateral triangle and the centre of mass of the system lies at it's centroid. The bodies revolve around the centre of mass in a symmetric orbit but after some period of time they are flung away from their seemingly stable orbit and chaotic motion takes over the system. Changing the initial conditions such that the bodies now lie at the vertices of a slightly bigger equilateral triangle produces the same type of orbit but the chaotic regime of this configuration is completely different from that of the initial configuration.

An unequal mass system that demonstrates similar behaviour is shown alongside (FIG. 9). The blue body has double the mass of other bodies. When the bodies approach the centre of mass of the system (which is slightly shifted towards blue), the other bodies go around in loops. This perturbation destabilises the system.

C. Chaotic systems

The systems corresponding to this classification are extremely chaotic in nature. Even the slightest change in the initial conditions ensures that the system traverses a path in the phase space that can never be replicated. Such systems are extremely susceptible to perturbations.

The model shown below (FIG. 10) demonstrates one of the defining features of chaotic three body systems, the reduction of the complexity of the system. Every system in the universe tries to attain stability and these systems are no different. They obtain stability by reducing from a three body system to a two body system by flinging away the third body.

Another system with completely different initial conditions, unequal masses and a planar restriction is shown below (FIG. 11). The outcome remains the same, in the sense that it's impossible to reproduce this system without taking into consideration every digit of the mantissa into subsequent calculations.

IV. CONCLUSION

Re-coding the Mandelbrot set to form the logistic map provided us with an explanation for why the bifurcations lined up with the real line and probed us to further study it's fractal nature. But a more interesting thing we found about the logistic map was it's chaotic behaviour at certain values of r and the periodically distributed islands

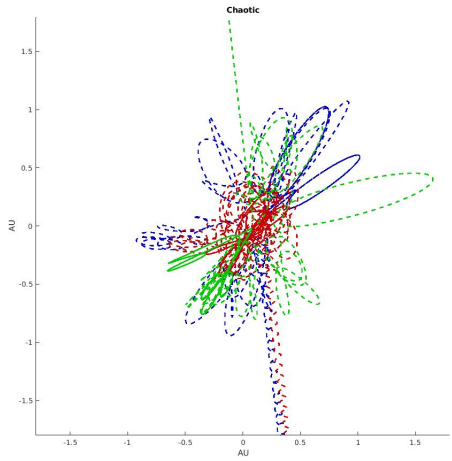


FIG. 10: Chaotic behaviour of an equal mass system

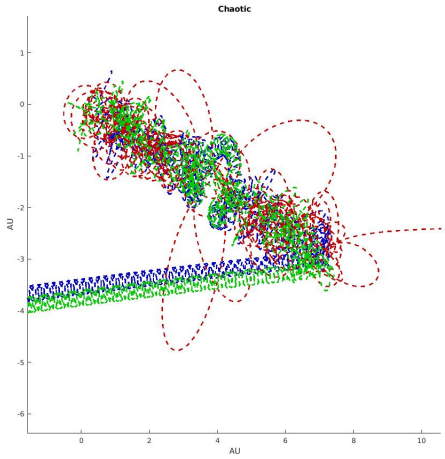


FIG. 11: Chaotic behaviour of an unequal mass system.

FIG. 12: The highlighting feature displayed in both systems is how the three body system reduces to a two body system to attain stability. The third body is always flung away with the help of the gravity assist of other two bodies

of stability, this lead us to probe deeper into chaos theory and the problems it posed. Lorenz demonstrated the onset of chaos and sensitive dependence on initial conditions in his experiment - the Lorenz water wheel. One of the widely studied application of chaos is the three body system and the N body system as a whole.

The three body problem lacked an analytical solution for centuries and Poincaré proved why it won't ever be solved analytically. In 1912, Karl Sundman found a series solution[12] for the problem but it converges so slowly that it is infeasible to produce practically relevant results from it. Since then the three body problem has been solved for several restrictive cases and the search is still going on for more viable and general systems. Classifying three body systems into long term stable, quasi-stable and chaotic systems helps to not only bring to light the difficulty in finding solutions for such systems but also it's complexity and the infinite solutions it has in certain regions of initial conditions. These can be equated to the islands of stability, only the problem here is that these islands are extremely scattered and thus give us a very small window of exploring them.

V. ACKNOWLEDGEMENTS

We would like to thank our mentor Andres Lopez Moreno for guiding and supporting us throughout the project. We would also like to thank ISEC for providing us this opportunity to be able to research and review topics on our own under expert guidance. I would also like to thank my colleague Shobhit Ranjan for proofreading the article.

[1] G.-J. Lay, *Verhulst-mandelbrot-bifurcation* (2008), URL `https://commons.wikimedia.org/wiki/File:Verhulst-Mandelbrot-Bifurcation.jpg`.

[2] M. J. Feigenbaum (1980), URL `https://permalink.lanl.gov/object/tr?what=info:lanl-repo/lareport/LA-UR-80-5007`.

[3] A. Provenzale and A. R. Osborne, in *Dynamics and Stochastic Processes Theory and Applications*, edited by R. Lima, L. Streit, and R. Vilela Mendes (Springer Berlin Heidelberg, Berlin, Heidelberg, 1990), pp. 260–275, ISBN 978-3-540-46969-8.

[4] E. N. Lorenz, *Deterministic Nonperiodic Flow* (Springer New York, New York, NY, 2004), pp. 25–36, ISBN 978-0-387-21830-4, URL `https://doi.org/10.1007/978-0-387-21830-4_2`.

[5] *Lorenz atrractor*, URL `https://matplotlib.org/3.1.0/_images/sphx_glr_lorenz_attractor_001.png`.

[6] C. A., *Poincare and the three-body problem* (2012), URL `http://www.bourbaphy.fr/chenciner.pdf`.

[7] V. Szebehely and K. Zare, **58**, 145 (1977), URL `https://ui.adsabs.harvard.edu/abs/1977A&A....58..145S`.

[8] S. A. Astakhov, A. D. Burbanks, S. Wiggins, and D. Farrelly, Nature (London) **423**, 264 (2003).

[9] D. V. Šuvakov M. (2013), URL `https://arxiv.org/abs/1303.0181`.

[10] V. D. M. Šuvakov, *Three classes of newtonian three-body*

planar periodic orbits (2013), URL http://three-body.ipb.ac.rs.

[11] M. R. Chenciner A. (2000), URL www.jstor.org/stable/2661357.

[12] S. P. (2013), URL https://sites.math.washington.edu/~morrow/336_13/papers/peter.pdf.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/343988891

Linear Feedback Shift Register-Based Test Pattern Generators - A Comparative Study

Article · August 2020

CITATIONS
0

READS
10

2 authors, including:



Parangat Mittal
Delhi Technological University
1 PUBLICATION **0** CITATIONS

SEE PROFILE

All content following this page was uploaded by Parangat Mittal on 31 August 2020.

The user has requested enhancement of the downloaded file.

ISSN: 2278 – 909X
International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 9, Issue 8, August 2020

Linear Feedback Shift Register-Based Test Pattern Generators: A Comparative Study

Parangat Mittal, Daksh Shah

Delhi Technological University, New Delhi

***Abstract*— Pseudo Random Number Generators are widely used in VLSI Design as Test Pattern Generators for testing of digital circuits in a BIST system. The test pattern sequence generated also finds applications in cryptography. It is thus essential to design an efficient test pattern generator which utilizes least hardware, dissipates lowest power and generates most random sequence. This paper attempts to compare a few recent proposals where Linear Feedback Shift Register (LFSR) has been suitably modified to obtain an efficient architecture of a test pattern generator.**

***Index Terms*—Test Pattern Generator (TPG), Linear Feedback Shift Register (LFSR), Testing of Digital Circuits, Switching Activity, Random Sequence**

I. INTRODUCTION

In the field of VLSI Design, Digital Circuits offer a number of advantages in terms of speed, area and power consumption over Analog Circuits. A lot of design methodologies and architectures are proposed in literature to improve these parameters. While improving circuit design parameters, it is also essential to increase the testability, i.e. the ability of fault detection and rectification of the circuit. Built-In Self-Test (BIST) is a such Design for Testability Technique aimed to improve testability of the Circuit Under Test (CUT), while ensuring low power dissipation. A typical BIST[1] architecture has four elements - Test Pattern Generator, Circuit Under Test, Output Response Analyzer and BIST Controller.

The Test Pattern Generator (TPG) produces test patterns that are applied as inputs to the CUT. Generally, Pseudo Random Number Generators (PRNGs) based on Linear Feedback Shift Register (LFSR) architecture are used for this purpose. Power dissipation becomes a major concern while applying these test patterns due to the high switching activity between consecutive patterns. As the power dissipation increases, the current drawn increases which might lead to damage of the digital circuit.

Previous works have suggested a number of architectures for TPGs. While some [2] have modified the existing LFSR with a combinational element, while some [3][4] have implemented the physical design with a different logic family. This paper attempts to do a comparative study on the former proposals, where LFSR has been modified or updated to get an improved Test Pattern Generator.

The parameters considered while comparing the architectures are hardware utilisation, i.e. no. of Flip Flops (FFs) and Look-Up Tables (LUTs), and measure of randomness of the test patterns.

II. LFSR ARCHITECTURE

A Linear Feedback Shift Register is a sequential circuit consisting of D Flip Flops and exclusive OR gates, arranged linearly. For a N-bit LFSR, N is referred to as the degree of the characteristic polynomial that defines its construction. The degree of the polynomial gives the number of Flip Flops, and the number of non-zero coefficients (excluding x_n and x_0) gives the number of XOR gates. The characteristic equation is of the form:

$$\varphi(x) = \varphi_n x_n + \varphi_{n-1} x_{n-1} \cdot \cdot \cdot + \varphi_1 x + \varphi_0$$

A. *Type -1 / Modular LFSR*

In an internal-XOR LFSR, the output of the last flip-flop in the shift register, D_{n-1}, is fed back to the inputs of a selected subset of flip-flops. In the characteristic equation, for i = 0, 1, . . ., n−1, $\varphi_i = 1$ indicates that feedback occurs at the input of D_i, and $\varphi_i = 0$ indicates that no feedback occurs there.

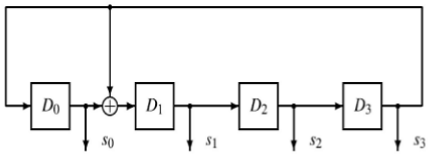


Figure 1 - Type - 1 LFSR

B. *Type - 2 / External LFSR*

In an external-XOR LFSR, the outputs of a selected subset of flip-flops are combined into a single signal, which is fed back to the input of the first flip-flop in the shift register, D_{n-1}. The fact that the signal is fed back to the input of D_{n-1} (which can be viewed as being driven by non-existent D_n) is indicated by always assigning $\varphi_n = 1$. For i = 0, 1, . . ., n−1, $\varphi_i = 1$ indicates that the output of D_i is selected for feedback. Once again, $\varphi_0 = \varphi_n = 1$ for a non-trivial n-stage external-XOR LFSR.

All Rights Reserved © 2012 IJARECE

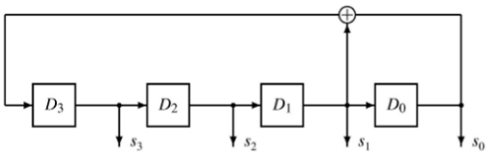


Figure 2 - Type - 2 LFSR

III. Comparative Study

A. *Low Power Test Pattern Generator* [5],[6]

This algorithm of test pattern generation aims to increase the correlation between test patterns to reduce the switching activity between primary inputs (PIs), which eventually scale down the power utilisation. This architecture consists of an LFSR with appended combinational logic to produce intermediary patterns with fewer transitions compared to the next one.

The algorithm consists of a Finite State Machine with 4 states to drive the test pattern generator. The LFSR is divided into two halves with an intermediary state-capturing Flip Flop in between. The half-LFSRs and output MUXs are enabled or disabled using control signals.

To begin with, the first half of the LFSR performs the shift operation, while the second half is disabled. The output is given by the LFSR output. Next, both the halves are idle, and the MUX performs the selective bit complement operation based on previous value of the output. The output comprises DFF outputs of the first half and bit-complemented output of the second.

Further, the first half of LFSR is disabled, while the second half performs the shift operation giving the output. At last, again both the halves are idle, and the MUX performs selective bit-complement based on previous value of the output. The output comprises bit-complemented results of the first half and DFF outputs of the second half.

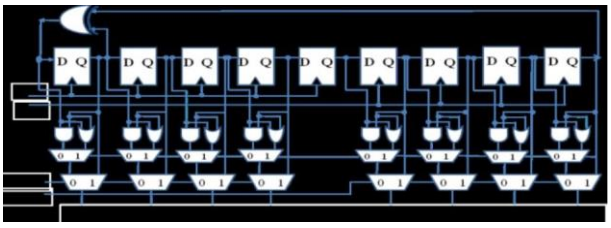


Figure 3 -Low Power LFSR [6]

B. *Bit Swapping LFSR* [7],[8]

This architecture aims to reduce the switching activity between consecutive outputs of the LFSR that serve as input to the CUT. The number of bit transitions in the subsequent outputs when minimised results in reduction of power as well. The LFSR is initiated with a non-zero seed polynomial.
The modification to a conventional LFSR is that one of the output lines (say N^{th}, for an N-bit LFSR) serves as a selection line for swapping the remaining bits. If that bit is 1 the bits are swapped, otherwise not. The bit swapping takes place in such a fashion that the 1st bit is swapped with 2nd bit, 3rd bit with 4th bit, and goes on till $(n-2)^{th}$ and $(n-1)^{th}$ bit (if n is odd), and till $(n-3)^{th}$ and $(n-2)^{th}$ bit (if n is even).

This bit swapping technique reduces the number of transitions between the subsequent patterns generated, and hence would result in lower power consumption during the testing phase of the CUT.

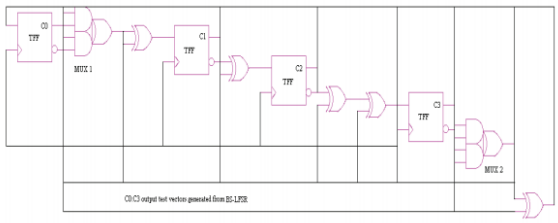


Figure 4 - Bit Swapping LFSR [7]

C. *Aperiodic LFSR* [9]

This architecture aims to introduce randomness among the patterns generated, and drift away from a fixed period after which the output of the test pattern generator starts repeating. It introduces control units for Feedback configuration Algorithm, Seed Configuration Algorithm and Range Configuration Algorithm to a traditional LFSR.

Feedback Configuration Algorithm:
The fixed feedback paths are broken by a control input, which is ANDed with every feedback path. If the value of this input is 0, feedback is disabled while on 1, it is enabled.

Seed Configuration Algorithm:
A random seed value is inserted at random intervals in the circuit. The seed comes from the samples of metastable state signals, which are almost unpredictable. For simulation purposes, the random function of Verilog was employed.

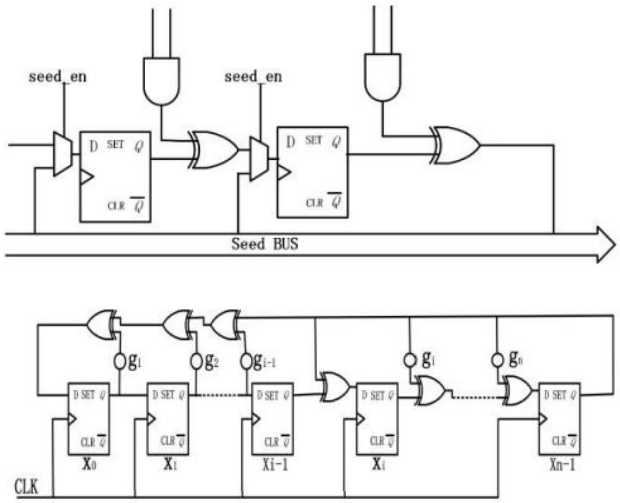


Figure 5 - Aperiodic LFSR [9]

IV. Methodology

All the LFSR architectures described in the previous two sections were designed in Verilog HDL. The circuit elements like 2-input logic gates, 2:1 Multiplexers, D Flip Flops etc. were implemented from standard Verilog libraries to

All Rights Reserved © 2020 IJARECE

maintain consistency. As a reference for uniformity, the following characteristic polynomial was chosen for all the LFSRs:

$$f(x) = 1 + x^4 + x^6 + x^{11} + x^{13} + x^{15}$$

The Verilog modules were simulated in Xilinx Vivado Design Suite 2018.3. The Elaborated Design (RTL) of each architecture has been attached in the above sections, where they were described. The Test Pattern Generators were tested against a common testbench with a 16-bit Ripple Carry Adder as the Circuit Under Test.

The 16-bit outputs of LFSR were tested for pseudo-randomness via NIST Statistical Test Suite[10]. The LSB's of every test vector were appended to form a single 100-bit sequence. The Frequency Monobit Test was performed in MATLAB, through a script.

V. Simulation Results

Table 1: Comparison of LFSR architectures

Architecture	Utilisation		
	FFs	LUTs	I/O
Conventional LFSR	32	23	19
Low Power LFSR	52	74	19
Bit Swapping LFSR	32	47	19
Aperiodic LFSR	32	39	21

Table 1 clearly shows that a conventional LFSR is the most efficient to implement, in terms of hardware utilisation. This is because no combinational circuit has been added to modify its functionality. While a Low Power LFSR having suitable modifications uses the most hardware.

Table 2: NIST Test Suite Results

Architecture	p-value	Type of sequence
Conventional LFSR	0.1615	Random
Low Power LFSR	0.6153	Random
Bit Swapping LFSR	0.5485	Random
Aperiodic LFSR	0.8132	Random

The Frequency Monobit Test of NIST Statistical Test Suite[10] states that a p value close to 1 indicates a truly random sequence and that close to 0 indicates non-randomness of a sequence. A p-value > 0 01 is generally considered as a random sequence. From Table 2, it can be stated that the Aperiodic LFSR generates the most random test sequence, while a conventional LFSR generates the least random sequence.

VI. Conclusion

This paper presented a comparative analysis of three modern LFSR-based Test Pattern Generators. While the design with the least hardware throughput, generates a test sequence which is less random than the one with a higher hardware utilisation. A low power design, as the name suggests, dissipates lowest power but is unable to generate a truly random sequence. Other designs that require fewer hardware components and generate truly random sequences can be studied and implemented in the future.

REFERENCES

[1] Babu, S. L., & Sarma, N. S. M. (2018). Automatic test pattern generation in VLSI - A survey. 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing, ICECDS 2017, 3012–3017.

[2] Saraswathi, T., Ragini, K., & Ganapathy Reddy, C. (2011). A review on power optimization of linear feedback shift register (LFSR) for Low power built in self test (BIST). ICECT 2011 - 2011 3rd International Conference on Electronics Computer Technology, 6, 172–176.

[3] Oommen, R., George, M. K., & Joseph, S. (2018). Study and Analysis of Various LFSR Architectures. 2018 International Conference on Circuits and Systems in Digital Enterprise Technology, ICCSDET 2018,1–6.

[4] Babu, S. L., & Sarma, N. S. M. (2018). Automatic test pattern generation in VLSI - A survey. 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing, ICECDS 2017, 3012–3017.

[5] Thubrikar, T., Kakde, S., Gaidhani, S., Kamble, S., & Shah, N. (2018). Design and implementation of low power test pattern generator using low transitions LFSR. Proceedings of the 2017 IEEE International Conference on Communication and Signal Processing, ICCSP 2017, 2018-January, 467–471.

[6] Saifuddin, K., Reddy, D. & Raviteja, C. (2019) Modified design of test pattern generator for built-in self test applications. International Journal for Innovative Engineering Management and Research, IJIEMR, Vol. 08, Issue 05, May 2019

[7] Dhanesh, P., & Jayanth Balaji, A. (2015). Dual threshold bit-swapping LFSR for power reduction in BIST. ICACCS 2015 - Proceedings of the 2nd International Conference on Advanced Computing and Communication Systems, 5–7.

[8] Abu-Issa, A.S., Quigley, S.F., Bit-swapping LFSR for low-power BIST (2008), Electronics Letters Vol. 44, no. 6, pp. 401-402, 13 March 2008.

[9] Tang, H., Qin, T., Hui, Z., Cheng, P., & Bai, W. (2018). Design and Implementation of a Configurable and Aperiodic Pseudo Random Number Generator in FPGA. 2018 IEEE 2nd International Conference on Circuits, System and Simulation, ICCSS 2018, 47–51.

[10] Rukhin, A., Soto, J., & Nechvatal, J. (2010). A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. Nist Special Publication, 22(April), 1/1--G/1.4

All Rights Reserved © 2012 IJARECE

View publication stats

arXiv:2008.11585v1 [cs.LO] 25 Aug 2020

LOGIC SIGNED PETRI NET

PAYAL AND SANGITA KANSAL

Department of Applied Mathematics, Delhi Technological University, Bawana Road,Delhi,India-110042
e-mail address: payal.dtu@gmail.com

Department of Applied Mathematics, Delhi Technological University, Bawana Road,Delhi,India-110042
e-mail address: sangita_kansal15@rediffmail.com

Abstract. In this paper,the authors show the versatility of the Signed Petri Net (SPN) introduced by them by showing the equivalence between a Logic Signed Petri Net (LSPN) and Logic Petri Net (LPN).The capacity of each place in all these nets is at most one, i.e., a place has either zero or one token in it.

1. Introduction

Petri Net is a well-known modeling tool to model concurrent systems.Since its inception, various extensions of PN's have been introduced viz., Continuous PN, Stochastic PN, Timed PN, Object PN, Hybrid PN, Workflow nets, Fuzzy PN, Lending PN, Multidimensional PN [AD98, Mar89, Mur89, Wan98, Val04, vdA98, CKC90, BCP15, vdA93] etc.The PN and its extensions have been widely used in various fields [HK18, FGS17, LYLT17, CFML17].To describe batch processing functions and passing value indeterminacy in cooperative systems, LPNs have been defined as high level PNs in [YG09].The authors introduced the concept of Signed Petri net (SPN)[KP] by utilizing the concept of signed graph and Petri net.SPNs are capable of modeling a large variety of systems and is prefered over a Petri net due to the presence of two types of tokens in it, positive and negative, which are distinguishable.Further, in comparison to a signed graph, SPN is advantageous since a single SPN can be used to represent various signed graphs by simply varying the marking of SPN due to firing of a sequence of transitions.Thus, we need to analyse one SPN in order to infer about all possible signed graph structures that can be formed for a fixed number of vertices.Although an SPN was introduced keeping in view the above mentioned advantages,in this paper, authors show the versatility of the SPN by showing that an equivalent LPN can be constructed using an SPN, which the authors call a Logic Signed Petri Net (LSPN).

ACM Classification:F.1.1.
Keywords: Signed Petri net, Petri net, marking, reachability tree.

© Payal and Sangita
Creative Commons

2. Basic Definitions

2.1. **Petri Net(**[Jen87]**).** A *Petri net (PN)* is a 5-tuple $N = (P, T, I^-, I^+, \mu_0)$,where

(1) P is the finite, non-empty set of places.
(2) T is the finite, non-empty set of transitions.
(3) $P \cap T = \emptyset$.
(4) $I^-, I^+ : (P \times T) \to \mathbb{N}$ where $\mathbb{N}$ is the set of non-negative integers, are called negative and positive incidence functions respectively.
(5) $\forall p \in P, \exists\ t \in T$ such that $I^-(p,t) \neq 0$ *or* $I^+(p,t) \neq 0$,and
$\forall t \in T, \exists\ p \in P$ such that $I^-(p,t) \neq 0$ *or* $I^+(p,t) \neq 0$
(6) $\mu_0 : P \to \mathbb{N}$ is the initial marking which gives the initial distribution of tokens in places.

The arc set of the Petri net N is defined as:

$$E = \{(p,t) : I^-(p,t) > 0\} \cup \{(t,p) : I^+(p,t) > 0\}$$

2.2. **Signed Petri Net(**[KP]**).**

Definition 2.1. Signed Petri Net (SPN)

A Signed Petri Net is defined as a 3-tuple $N^* = (N', \sigma, \mu_0)$,where

(1) $N' = (P, T, I^-, I^+)$ is a Petri net structure.
(2) $\sigma : E \to \{+, -\}$, where E is the arc set of N'.An arc is called a positive or negative arc respectively according to the sign + or − assigned to it using the function σ.
(3) $\mu_0 = (\mu_0^+, \mu_0^-)$ is the initial marking of SPN where
   (a) $\mu_0^+ : P \to \mathbb{N}$ gives the initial distribution of positive tokens in the places,called positive marking of SPN.
   (b) $\mu_0^- : P \to \mathbb{N}$ gives the initial distribution of negative tokens in the places,called negative marking of SPN.

Thus,a marking in SPN can be represented as a vector $\mu = (\mu^+, \mu^-)$ with $\mu^+, \mu^- \in \mathbb{N}^n, n = |P|$ such that $\mu(p_i) = (\mu^+(p_i), \mu^-(p_i))\ \forall\ p_i \in P$.

Graphically, positive and negative arcs in an SPN are represented by solid and dotted lines respectively.A positive token is represented by a filled circle and a negative token by an open circle.
An SPN is said to be *negative* if all of its arcs are negative in sign.

Remark 2.2. $N'' = (N', \sigma)$ is called an SPN structure where N' is a PN structure and $\sigma : E \to \{+, -\}$.

2.2.1. *Execution Rules for Signed Petri Net.* Similar to a Petri net, the execution of an SPN depends on the distribution of tokens in its places.The execution takes place by firing of a transition.A transition may fire if it is enabled.
A transition t in an SPN N^* is *enabled* at a marking $\mu = (\mu^+, \mu^-)$ if

$$I^-(p,t) \leq \mu^+(p)\ \forall p \in {}^\bullet t \text{ for which } \sigma(p,t) = +$$
$$I^-(p,t) \leq \mu^-(p)\ \forall p \in {}^\bullet t \text{ for which } \sigma(p,t) = -$$

An enabled transition t may *fire* at $\mu = (\mu^+, \mu^-)$ provided $\exists p_k \in t^\bullet$ such that:

$$\sigma(t, p_k) = \begin{cases} + & \text{if } \sigma(p,t) = + \ \forall p \in {}^\bullet t \\ - & \text{if } \sigma(p,t) = - \ \forall p \in {}^\bullet t \\ + \text{ or } - & \text{if } \sigma(p,t) = + \text{ for some } p \in {}^\bullet t \ \textit{and} \ - \text{ for some } p \in {}^\bullet t \end{cases}$$

After firing, it yields a new marking $\mu_1 = (\mu_1^+, \mu_1^-)$ given by the rule:

$$\mu_1^+(p) = \mu^+(p) - I^-(p,t) + I^+(p,t) \ \forall p \in P \quad \text{where (p,t) \& (t,p) are positive arcs, if exist}$$

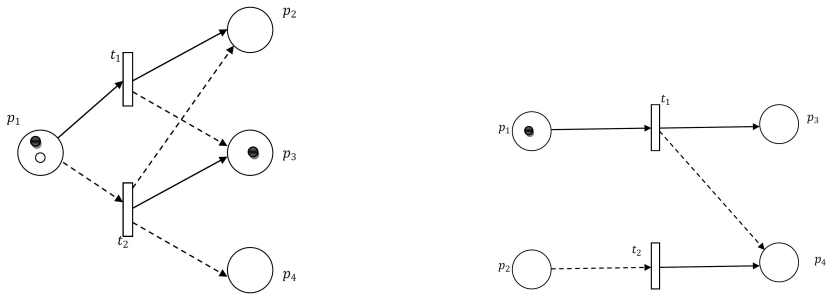
$$\mu_1^-(p) = \mu^-(p) - I^-(p,t) + I^+(p,t) \ \forall p \in P \quad \text{where (p,t) \& (t,p) are negative arcs, if exist}$$

We say that μ_1 is reachable from μ and write $\mu \xrightarrow{t} \mu_1$.We restrict the movement of positive(negative) tokens to positive(negative) arcs only.

A marking μ is reachable from μ_0 if there exists a firing sequence η that transforms μ_0 to μ and is written $\mu_0 \xrightarrow{\eta} \mu$.A *firing or occurence sequence* is a sequence of transitions $\eta = t_1 t_2 \dots t_k$ such that

$$\mu_0 \xrightarrow{t_1} \mu_1 \xrightarrow{t_2} \mu_2 \xrightarrow{t_3} \mu_3 \dots \xrightarrow{t_k} \mu$$

Note that a transition $t_j, 1 \leq j \leq k$ can occur more than once in the firing sequence η.
Let us look at the execution of an SPN with the help of an example.
In figure $1(a)$, t_1 and t_2 both are enabled at μ_0. Firing of t_1 yields a new marking $\mu = ((0,1,1,0),(1,0,1,0))$ and firing of t_2 yields a new marking $\mu = ((1,0,2,0),(0,1,0,1))$. In figure $1(b)$, t_1 is enabled,while t_2 is not. t_1 can fire to give a new marking $\mu = ((0,0,1,0),(0,0,0,1))$.



(A) SPN with $\mu_0 = ((1,0,1,0),(1,0,0,0))$

(B) SPN with $\mu_0 = ((1,0,0,0),(0,0,0,0))$

FIGURE 1. Execution of an SPN

Definition 2.3. Reachability Set of Signed Petri net

The Reachability Set $R(N^*, \mu_0)$ of an SPN N^* is the set of all markings of N^* reachable from the initial marking μ_0.

Definition 2.4. Logic Petri net([YG09])

$LPN = (P, T, I^-, I^+, I, O, \mu_0)$ is a logic PN if and only if

(1) (P, T, I^-, I^+, μ_0) is a PN.

(2) T includes three subsets of transitions,i.e.,$T = T_D \cup T_I \cup T_O$, $\forall\, t \in T_I \cup T_O : {}^\bullet t \cap t^\bullet = \emptyset$ where
- T_D denotes a set of traditional transitions.
- T_I denotes a set of logic input transitions.
- T_O denotes a set of logic output transitions.

(3) I is a mapping from logic input transitions to a logic input expression,i.e.,$\forall\ t \in T_I, I(t) = f_I(t)$, where $f_I(t)$ is a logic input expression associated with the transition t.

(4) O is a mapping from logic output transitions to a logic output expression,i.e.,$\forall\ t \in T_O, O(t) = f_O(t)$, where $f_O(t)$ is a logic output expression associated with the transition t.

(5) *Transition Firing Rules*
- $\forall t \in T_D$,the firing rules are same as in PNs.
- $\forall t \in T_I$, t is enabled at μ if the input expression $f_I(t)$ is true at μ.After firing it yields a new marking μ' given by the following rule:
  $\forall p \in {}^\bullet t$, if $\mu(p) = 1$, then $\mu'(p) = \mu(p) - 1$;
  $\forall p \in t^\bullet, \mu'(p) = \mu(p) + 1$
  $\forall p \notin {}^\bullet t \cup t^\bullet, \mu'(p) = \mu(p)$
- $\forall t \in T_O$,t is enabled if $\forall p \in {}^\bullet t : \mu(p) = 1$. Firing t will generate a new marking μ', given by: $\forall p \in S : \mu'(p) = \mu(p) + 1$;
  $\forall p \in {}^\bullet t, \mu'(p) = \mu(p) - 1$
  $\forall p \notin {}^\bullet t \cup t^\bullet$ or $\forall p \in t^\bullet \backslash S : \mu'(p) = \mu(p)$ where $S \subseteq t^\bullet$, such that output expression $f_O(t)$ is true at μ'.

Remark 2.5. In the definition of Logic Petri Net, a logic input expresssion is attached to a logic input transition.By this we mean a logic expression involving the input places of the given transition.
Similarly, a logic output expresssion is attached to a logic output transition.By this we mean a logic expression involving the output places of the given transition.

3. Logic Signed Petri Net(LSPN)

The Logic Signed Petri Net (LSPN) has been obtained by modification of execution rules of an SPN as defined in [KP].

Definition 3.1. Logic Signed Petri net

A logic SPN is defined as an SPN $N^* = (P, T, A^-, A^+, B^-, B^+, \mu_0)$ where

(i) P is a finite,non empty set of places.

(ii) T includes three subsets of transitions,i.e.,$T = T_D \cup T_I \cup T_O$, $\forall\, t \in T_I \cup T_O : {}^\bullet t \cap t^\bullet = \emptyset$ where
- T_D denotes a set of traditional transitions.
- T_I denotes a set of logic input transitions.
- T_O denotes a set of logic output transitions.

If $t \in T_I$ then $\forall\ p\ \in {}^\bullet t$ either $\sigma(p,t) = +$ or $\sigma(p,t) = +$ as well as $-$ and $\forall\ p\ \in t^\bullet$ $\sigma(t,p) = +$.
Similarly, if $t \in T_O$ then $\forall\ p\ \in t^\bullet$ either $\sigma(t,p) = +$ or $\sigma(t,p) = +$ as well as $-$ and $\forall\ p\ \in {}^\bullet t\ \sigma(p,t) = +$.

(iii) $A^- = [a_{ij}^-]$ where a_{ij}^- gives the number of positive arcs from p_j to t_i.
(iv) $A^+ = [a_{ij}^+]$ where a_{ij}^+ gives the number of positive arcs from t_i to p_j.
(v) $B^- = [b_{ij}^-]$ where b_{ij}^- gives the number of negative arcs from p_j to t_i.
(vi) $B^+ = [b_{ij}^+]$ where b_{ij}^+ gives the number of negative arcs from t_i to p_j.
(vii) A marking in an LSPN can be represented as a vector $\mu = (\mu^+, \mu^-)$ with $\mu^+, \mu^- \in \mathbb{N}^n, n = |P|$ such that $\mu(p_i) = (\mu^+(p_i), \mu^-(p_i))\ \forall\ p_i \in P$.

3.1. **Execution rules for LSPN.**

(i) If $t_i \in T_D$, then the execution rules are same as for an SPN.
(ii) For $t_i \in T_I$

- *Enabling Condition-* A transition $t_i \in T_I$ is enabled at μ provided
  (i) $\mu^+(p_j) = 1\ \forall\ p_j\ \in\ S_1 = \{p_j \in {}^\bullet t_i \mid a_{ij}^- = 1, b_{ij}^- = 0\}$ and
  (ii) $\forall p\ \in\ S_2 = \{p_j \in {}^\bullet t_i \mid a_{ij}^- = 1, b_{ij}^- = 1\}$, $\exists p_k \in S_2$ with $\mu^+(p_k) = 1$ and $\forall p_j \in S_2 \backslash \{p_k\}$, either $\mu^+(p_j) = 1$ or $\mu^-(p_j) = 1$.
- *Firing Condition-* An enabled transition $t_i \in T_I$ may fire at a marking $\mu = (\mu^+, \mu^-)$ to yield a new marking μ_1 given by :-

  – $\forall\ p_j\ \in\ S_1$,
  $$\mu_1^+(p_j) = \mu^+(p_j) - a_{ij}^- + a_{ij}^+$$
  – $\forall\ p_j\ \in\ S_2$, for which $\mu^+(p_j) = 0$ & $\mu^-(p_j) = 1$
  $$\mu_1^+(p_j) = \mu^+(p_j)$$
  $$\mu_1^-(p_j) = \mu^-(p_j) - b_{ij}^- + b_{ij}^+$$
  – $\forall\ p_j\ \in\ S_2$, for which $\mu^+(p_j) = 1$ & $\mu^-(p_j) = 0$
  $$\mu_1^-(p_j) = \mu^-(p_j)$$
  $$\mu_1^+(p_j) = \mu^+(p_j) - a_{ij}^- + a_{ij}^+$$
  – For all $p_j \in P \backslash {}^\bullet t_i$
  $$\mu_1^+(p_j) = \mu_0^+(p_j) - a_{ij}^- + a_{ij}^+$$
  $$\mu_1^-(p_j) = \mu_0^-(p_j) - b_{ij}^- + b_{ij}^+$$

(iii) For $t_i \in T_O$

- *Enabling Condition-* A transition $t_i\ \in T_O$ is enabled if $\forall p_j\ \in\ S_1, \mu^+(p_j) = 1$.
- *Firing Condition-*An enabled transition $t_i \in T_O$ may fire at a marking $\mu = (\mu^+, \mu^-)$ to yield a new marking μ_1 given by :-
  $\forall p_j \in t_i^\bullet$
  – $\mu_1^+(p_j) = 1$ whenever $a_{ij}^+ = 1, b_{ij}^+ = 0$
  – If $a_{ij}^+ = 1, b_{ij}^+ = 1$,then,$\exists p_k \in t_i^\bullet$ with $\mu_1^+(p_k) = 1$ and $\forall p_l \in t_i^\bullet \backslash \{p_k\}$ either $\mu_1^+(p_l) = 1$ or $\mu_1^-(p_l) = 1$

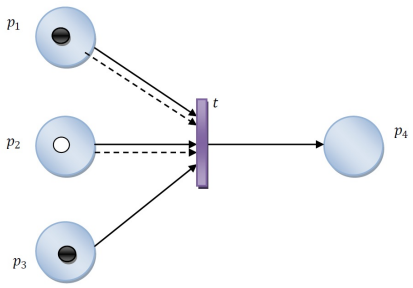


FIGURE 2. A LSPN with a logic input transition.Transition t is enabled at $\mu = ((1,0,1,0),(0,1,0,0))$. t fires to yield a new marking $\mu_1 = ((0,0,0,1),(0,0,0,0))$

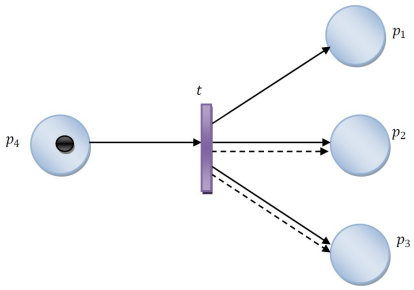


FIGURE 3. A LSPN with a logic output transition.Transition t is enabled at $\mu = ((0,0,0,1),(0,0,0,0))$. t fires to yield a new marking $\mu_1 \in ((1,1,0,0),(0,0,1,0)),((1,0,1,0),(0,1,0,0)),((1,1,1,0),(0,0,0,0))$

$$\forall p_j \in P \backslash t_i^{\bullet}$$
$$\mu_1^+(p_j) = \mu_0^+(p_j) - a_{ij}^- + a_{ij}^+$$
$$\mu_1^-(p_j) = \mu_0^-(p_j) - b_{ij}^- + b_{ij}^+$$

While modeling a workflow process using an LSPN, a positive token in a data place represents that the data has arrived from the organization representing this place, while a negative token represents that the data has not arrived from the organization.So, in a given workflow cycle, all the arrived data is processed while the data which arrives late is processed in the next workflow cycle.

4. Equivalence of LPN and LSPN

In order to prove the equivalence between LPN and LSPN, their isomorphism and equivalent definitions are given.

Definition 4.1. Isomorphism
Let $N' = (P, T, I^-, I^+, I, O, \mu_0)$ be an LPN and $N^* = (P', T', A^-, A^+, B^-, B^+, \mu_0')$ be an LSPN.Let $R(N', \mu_0)$ be the reachability set of N' and , $R^+(N^*, \mu_0')$ be the reachability set of positive markings of N^*.

Then, $R(N', \mu_0)$ & $R^+(N^*, \mu_0')$ are isomorphic if there exists a bijective function $f : R(N', \mu_0) \to R^+(N^*, \mu_0')$such that

$$\forall \mu_1, \mu_2 \in R(N', \mu_0), t \in T; \mu_1 \xrightarrow{t} \mu_2 \implies \exists t' \in T'; f(\mu_1) \xrightarrow{t'} f(\mu_2)$$

Definition 4.2. Equivalence
Let $N' = (P, T, I^-, I^+, I, O, \mu_0)$ be an LPN and $N^* = (P', T', A^-, A^+, B^-, B^+, \mu_0')$ be an LSPN.Then, N' & N^* are equivalent iff $R(N', \mu_0)$ & $R^+(N^*, \mu_0')$ are isomorphic to each other.

Theorem 4.3. *For any Logic Petri Net(LPN) there exists an equivalent Logic Signed Petri Net (LSPN).*

Proof. Given an LPN $N' = (P, T, I^-, I^+, I, O, \mu_0)$, we construct an LSPN $N^* = (P', T', A^-, A^+, B^-, B^+, \mu_1)$ as follows.

(1) The set of places of N^* is same as N' ,i.e. $P' = P$.
(2) Construction of set of transitions T' of N^*:
   (i) All the traditional transitions of N' will be the traditional transitions of N^* i.e. if $t_i \in T_D = T \backslash (T_I \cup T_O)$.Then $t_i \in T_D' \subseteq T'$.This traditional transition t_i in N^* is connected with a place in N^* given by the rule below:

$$\forall p_j \in P, \text{ if } (p_j, t_i) \in I^- \text{ then } a_{ij}^- = 1, b_{ij}^- = 0$$

$$\forall p_k \in P, \text{ if } (p_k, t_i) \in I^+ \text{ then } a_{ik}^+ = 1, b_{ik}^+ = 0$$

   (ii) All the logic input transitions of N' are logic input transitions of N^* i.e. if $t_i \in T_I = T \backslash (T_D \cup T_O)$.Then $t_i \in T_I' \subseteq T'$ and this transition $t_i \in T_I'$ is connected with a place in N^* by the rule below:
   Let ${}^\bullet t_i = \{p_1, p_2, p_3, ..., p_k\}$ in N' and $f_I(t_i)$ be the logic input expression associated with t_i, convert $f_I(t_i)$ into its disjunctive normal form (DNF), which is unique.If any p_j occurs in both the forms i.e. p_j and $\neg p_j$ in the DNF then there exist two arcs in N^* from p_j to t_i, one of positive sign and other of negative sign.Therefore, for such $p_j \in {}^\bullet t_i,\ a_{ij}^- = 1, b_{ij}^- = 1$.
   On the other hand, if p_j occurs only in positive form i.e. as p_j and not as $\neg p_j$ in the DNF, then a positive arc is formed from p_j to t_i. Therefore, for such $p_j \in {}^\bullet t_i,\ a_{ij}^- = 1, b_{ij}^- = 0$. $\forall p_k \in t_i^\bullet,\ a_{ik}^+ = 1, b_{ik}^+ = 0$.
   (iii) All the logic output transitions of N' are logic output transitions of N^* i.e. if $t_i \in T_O = T \backslash (T_D \cup T_I)$.Then $t_i \in T_O' \subseteq T'$ and this transition $t_i \in T_O'$ is connected with a place in N^* by the rule below:
   Let $t_i^\bullet = \{p_1, p_2, p_3, ..., p_k\}$ in N' and $f_O(t_i)$ be the logic output expression associated with t_i, convert $f_O(t_i)$ into its disjunctive normal form (DNF) which is unique.If any p_j occurs in both the forms i.e. p_j and $\neg p_j$ in the DNF then there exist two arcs in N^* from t_i to p_j, one of positive sign and other of negative sign.Therefore, for such $p_j \in t_i^\bullet,\ a_{ij}^+ = 1, b_{ij}^+ = 1$.
   On the other hand, if p_j occurs only in positive form i.e. as p_j and not as $\neg p_j$ in the DNF, then a positive arc is formed from t_i to p_j . Therefore, for such

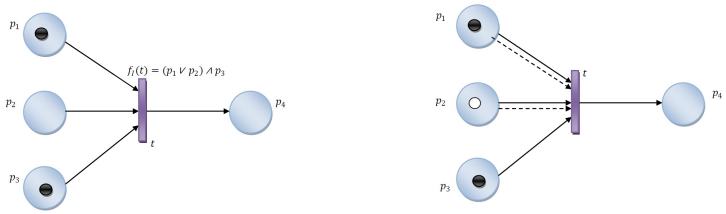
$p_j \in t_i^{\bullet},\ a_{ik}^{+} = 1, b_{ik}^{+} = 0.$
$\forall p_k \in {}^{\bullet}t_i,\ a_{ik}^{-} = 1, b_{ik}^{-} = 0.$

(3) Assignment of tokens
   (a) For all p_j satisfying ($a_{ij}^{-} = 1$ and $b_{ij}^{-} = 1$) or ($a_{ij}^{+} = 1$ and $b_{ij}^{+} = 1$)
   $\mu_1^{+}(p_j) = 1$ & $\mu_1^{-}(p_j) = 0$ in N^* if $\mu_0(p_j) = 1$ in N' and
   $\mu_1^{+}(p_j) = 0$ & $\mu_1^{-}(p_j) = 1$ in N^* if $\mu_0(p_j) = 0$ in N'
   (b) For all p_j which has not been assigned a token in step (a) and satisfying either $(a_{ij}^{-} = 1, b_{ij}^{-} = 0)$ or $(a_{ij}^{+} = 1, b_{ij}^{+} = 0)$,
   $\mu_1^{+}(p_j) = 1$ & $\mu_1^{-}(p_j) = 0$ in N^* if $\mu_0(p_j) = 1$ in N' and

   $\mu_1^{+}(p_j) = 0$ & $\mu_1^{-}(p_j) = 0$ in N^* if $\mu_0(p_j) = 0$ in N'

The equivalence between N' and N^* can be easily proved, because each marking of N' corresponds to a positive marking in N^* with the same number of positive tokens in the corresponding places in LPN and LSPN (By Construction).That is, in N' $\forall \mu_1, \mu_2 \in R(N', \mu_0), t \in T; \mu_1 \xrightarrow{t} \mu_2 \implies \exists t' \in T'; f(\mu_1) \xrightarrow{t'} f(\mu_2)$

This means that N' and N^* have same behaviour characteristics.Moreover, the structure of N^* is unique since the DNF is unique.So, f is a bijective function and $R(N', \mu_0)$ & $R^{+}(N^*, \mu_1)$ are isomorphic. Consequently, N' and N^* are equivalent. □



(A) LPN with a logic input transition

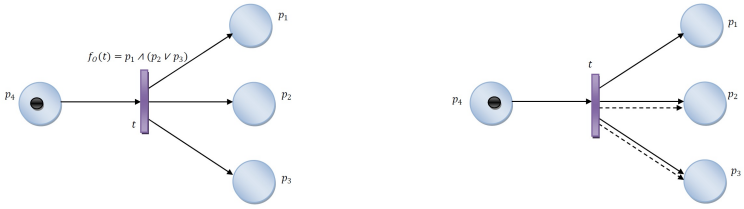
(B) Equivalent LSPN for LPN (A)

FIGURE 4. LPN and Equivalent LSPN

In the figure 4, an LPN with a logic input transition is converted to an equivalent LSPN using the procedure in the above theorem.

In the figure 5, an LPN with a logic output transition is converted to an equivalent LSPN using the procedure in the above theorem.

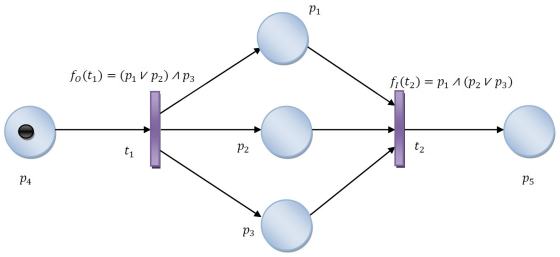
In the example in figure 6 , an LSPN is constructed from a given LPN using the procedure given in theorem above.The reachability tree of LPN and LSPN are also given in figure 7 and it can be clearly seen that corresponding to every marking in LPN we have a positive marking in LSPN.Thus, the reachability set of LPN and the set of positive markings of LSPN are isomorphic which implies that the two PNs are equivalent.



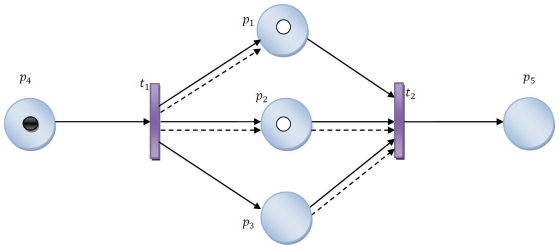
(A) LPN with a logic output transition

(B) Equivalent LSPN for LPN (A)

FIGURE 5. LPN and Equivalent LSPN



(A) AN LPN



(B) Equivalent LSPN

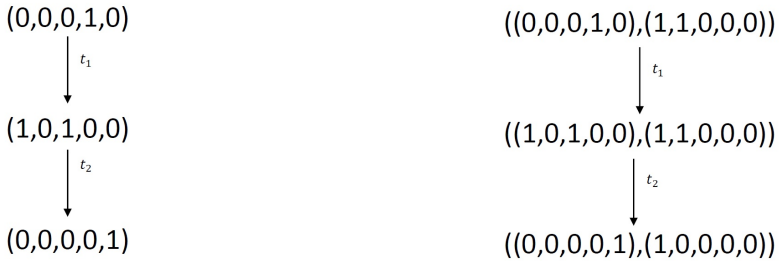
FIGURE 6. Example

5. CONCLUSION

An LSPN formulated using SPN by mere change of firing rules of the SPN can simulate an LPN.This shows the versatile nature of the SPN introduced by the authors.

REFERENCES

[AD98] H. Alla and R. David. Continuous and Hybrid Petri Nets. *Journal of Circuits,Systems and Computers*, 8, No. 1:159–188, 1998.



(A) Reachability Tree for LPN

(B) Reachability Tree for LSPN

FIGURE 7. Reachability Trees for LSPN Example in Figure 6

[BCP15] M. Bartoletti, T. Cimoli, and G.M. Pinna. Lending Petri Nets. *Science of Computer Programming*, pages 75–101, 2015.

[CFML17] Xuya Cong, Maria Pia Fanti, Agostino Marcello Mangini, and Zhiwu Li. Decentralized fault diagnosis by petri nets and integer linear programming. *IFAC-PapersOnLine*, 50(1):13624 – 13629, 2017. 20th IFAC World Congress.

[CKC90] Shyi-Ming Chen, Jyh-Sheng Ke, and Jin-Fu Chang. Knowledge representation using fuzzy petri nets. *IEEE Transactions on Knowledge and Data Engineering*, 2(3):299–309, 1990.

[FGS17] Angela Di Febbraro, Davide Giglio, and Nicola Sacco. On analyzing the vulnerabilities of a railway network with petri nets. *Transportation Research Procedia*, 27:553 – 560, 2017. 20th EURO Working Group on Transportation Meeting, EWGT 2017, 4-6 September 2017, Budapest, Hungary.

[HK18] Loic Helouet and Karim Kecir. Realizability of schedules by stochastic time petri nets with blocking semantics. *Science of Computer Programming*, 157:71 – 102, 2018. Special issue from PETRI NET 2016.

[Jen87] K. Jensen. Coloured petri nets. *LNCS,Springer, Heidelberg*, 254:248–299, 1987.

[KP] S. Kansal and Payal. An introduction to Signed Petri Net. *Communicated*.

[LYLT17] Hu-Chen Liu, Jian-Xin You, ZhiWu Li, and Guangdong Tian. Fuzzy petri nets for knowledge representation and reasoning: A literature review. *Engineering Applications of Artificial Intelligence*, 60:45 – 56, 2017.

[Mar89] M.A. Marsan. Stochastic petri nets:an elementary introduction. *LNCS ,Advances in Petri Nets,Springer*, pages 1–30, 1989.

[Mur89] T. Murata. Petri nets: Properties, analysis and applications. *Proceedings of IEEE*, 77(4):541–580, 1989.

[Val04] R. Valk. Object petri nets. *LNCS 3098,Lecture on Concurrency and Petri Nets,Springer*, pages 819–848, 2004.

[vdA93] W. M. P. van der Aalst. Multi-dimensional Petri nets. *Computing science notes Eindhoven:Technische Universiteit Eindhoven*, 9326, 1993.

[vdA98] W. M. P. van der Aalst. The application of Petri nets to workflow management. *Journal of Circuits, Systems and Computers*, 8(1):21–66, 1998.

[Wan98] J. Wang. Time Petri nets. In: Timed Petri Nets. *The Kluwer International Series on Discrete Event Dynamic Systems,Springer,Boston MA*, 9:63–123, 1998.

[YG09] Y. Y.Du and B.Q. Guo. Logic Petri Nets and Equivalency. *Information Technology Journal*, pages 95–100, 2009.

Low Power Quaternary Adder Using CNFET

Anurag Chauhan
Department of ECE
Delhi Technological University
Delhi, India
anurag2305chauhan@gmail.com

Lokesh Mahor
Department of ECE
Delhi Technological University
Delhi, India
lokesh.mahor15@gmail.com

Piyush Tiwari
Department of ECE
Delhi Technological University
Delhi, India
piyushtewari127@gmail.com

***Abstract*— Binary logic and MOS devices have been in use since the inception of the design era, but now due to advancement in VLSI industry binary logic has become tedious and complicated. To overcome this challenge Multi-valued logic (MVL) such as ternary and Quaternary Logic (QTL) can be used. MVL designs have an advantage over binary logic designs with respect to the chip area and the complexity of interconnection. In this paper, we present the design and performance analysis of QTL Full Adder (QFA) using Carbon Nano-tube Field Effect Transistor (CNFET). For design purposes, we have used the Stanford Virtual-Source Carbon Nanotube Field Effect Transistor Model version 1.01 with sub 10nm CNFET technology. The design tool used for simulation is Cadence Virtuoso. The proposed QFA design has been compared against the existing CNFET based QTL designs and it is found that the proposed QFA design is *97-98%* better in terms of Power Delay Product (PDP) and Energy Delay Product.**

Keywords*— *Multi-Value Logic, Quaternary Logic, CNFET, Adder.

I. INTRODUCTION

Over the last few decades, semiconductor devices have been a vital part of the electronics industry. As MOSFET continues to extend into the nanoscale regime, various problems have been observed which affect the electrical properties of the gate-oxide, threshold voltage (V_{th}) and saturation velocity (V_{sat}). This eventually introduces Short Channel Effects (SCE) such as hot electron effect, minimization of Drain Induced Barrier Lowering (DIBL), Impact-ionization and Surface Scattering [1], [2]. Due to these reasons, the necessity of new devices such as Double Gate Field-Effect Transistors (DGFET), Fin Field Effect Transistor (FinFET) , Tunnel Field-Effect Transistor (TFET) and Carbon Nanotube Field-Effect Transistors (CNFET) have been introduced [3].

Another complexity arises due to the scaling down of the technology as it results in an increase in the number of interconnections. As the technology node decreases the number of transistors increase, which causes the problem in routing of interconnects over the chip [4]. Using MVL reduces the chip area by reducing the interconnections and functional units in the VLSI design [5]. In MVL the basic principle is to send a large amount of data over a single wire, for example for transmitting the binary logic (00, 01, 10, 11) we need two wires, instead, in MVL (0, 1, 2, 3) we can transmit the same information over a single wire thus reducing the interconnection complexity[5].

Several QTL adder designs are proposed in literature[4], [6]–[8]. Adder design presented in [6], [7], [9] are decoder based designs and these designs have major drawback that they require MVL to be first converted in binary. All computation is done in binary and then the result is converted back into MVL. That's why these designs have high PDP and EDP. In this paper multiplexer-based QFA is implemented based on novel multiplexer design presented in [4].

The rest of the paper is organized as follows: Introduction to QTL basic gates implementation is given in section II. The proposed designs are presented in section III. Performance analysis for the proposed QFA design is given in section IV. Section V presents the conclusion.

II. QUATERNARY LOGIC

Quaternary Logic (QTL) is a base-3 number system and has logic levels 0, 1, 2 and 3 equivalents to 0 V, 0.2 V, 0.4 V, and 0.6 V respectively. QTL has four basic types of inverters and their logical definitions are given as follows:

$$SQI = 3 - x \tag{1}$$

$$PQI = \begin{cases} 3 & \text{if } x \neq 3 \\ 0 & \text{if } x = 3 \end{cases} \tag{2}$$

$$NQI = \begin{cases} 0 & \text{if } x \neq 0 \\ 3 & \text{if } x = 0 \end{cases} \tag{3}$$

$$IQI = \begin{cases} 3 & \text{if } x = 0,1 \\ 0 & \text{if } x = 2,3 \end{cases} \tag{4}$$

Table I presents the truth table for the QTL inverters. Fig. 1 shows the CNFET implementation of SQI, NQI, PQI and IQI inverters [7]. In this implementation we used Stanford Virtual-Source Carbon Nanotube Field Effect Transistor Model version 1.01 with sub 10nm CNFET technology [10], [11].

TABLE I. THE TRUTH TABLE FOR QTL INVERTERS

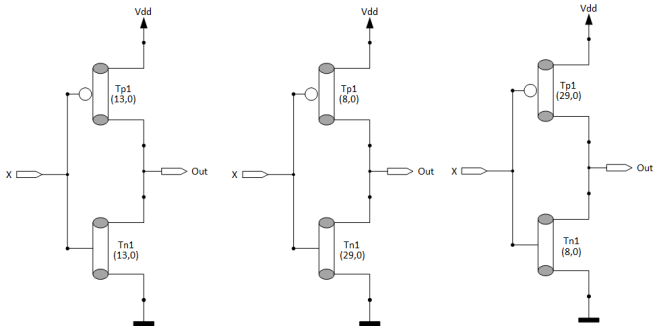
Input (x)	Output(s)			
	NQI	***IQI***	***PQI***	***SQI***
0	3	3	3	3
1	0	3	3	2
2	0	0	3	1
3	0	0	0	0

Fig. 2 (a) shows the CNFET implementation of a 4:1 multiplexer [4]. It consists of 6 inverters NQI, PQI, IQI, and 3 SQIs in order to obtain the signals *NQI, PQI, IQI,* $\overline{NQI}$, $\overline{PQI}$ and $\overline{IQI}$. These 6 signals are used to control the 6 Transmission gates, T_1 to T_6, which consist of NCNFET and PCNFET using CNT (29,0). Positive QTL 2:1 Multiplexer

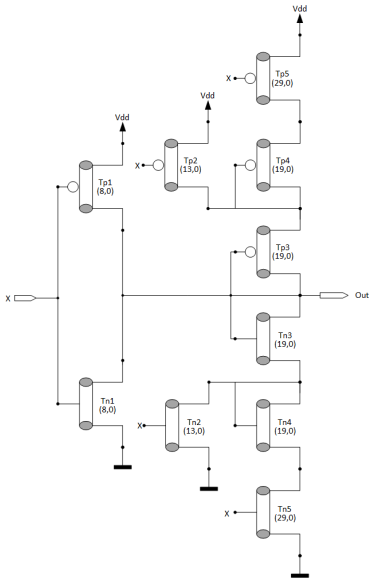
978-1-7281-1933-5/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:27:03 UTC from IEEE Xplore. Restrictions apply.

design is shown in Fig. 2 (b) [4]. It consists of inverter NQI and transmission gates T_1 and T_2.

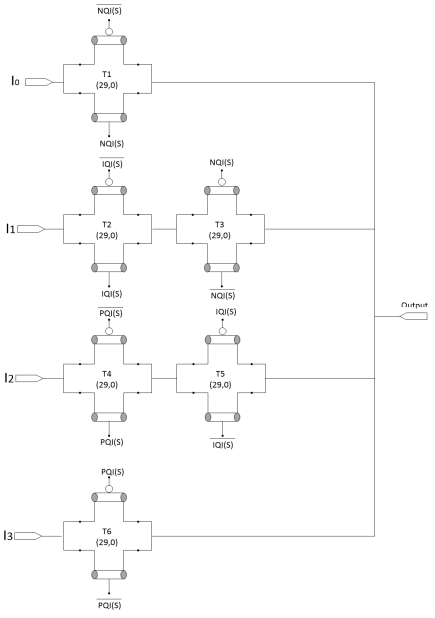


(a) (b) (c)

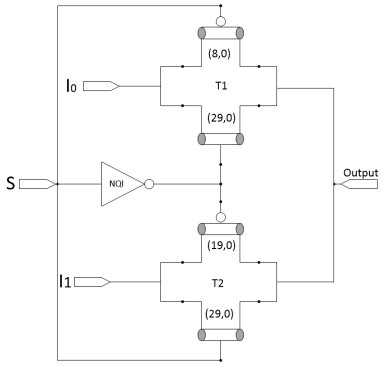


(d)

Fig. 1. CNFET implementation of (a) IQI, (b) NQI, (c) PQI and (d) SQI [3]



(a)



(b)

Fig. 2. CNFET implementation of (a) 4:1 multiplexer and (b) 2:1 multiplexer

III. Proposed Designs

A. Proposed QTL Functions

Modified 4:1 multiplexer is used to implement QTL functions incrementor, level-II incrementor and decrementor that is used in designing the proposed QFA design. In this implementation, we have only used either P-CNFET or N-CNFET instead of transmission gate based on the value of the input. The truth table for incrementor, level II incrementor, and decrementor are shown in Table II.

TABLE II. The truth table for incrementor, Level-II incrementor, and decrementor

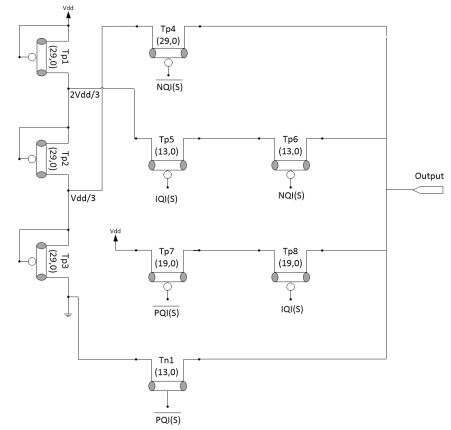
Input (x)	Output(s)		
	Incrementor	*Level II Incrementor*	*Decrementor*
0	1	2	3
1	2	3	0
2	3	0	1
3	0	1	2

Fig. 3 (a) shows the CNFET implementation of QTL incrementor. In this implementation Tp1, Tp2, and Tp3 are used to implement the voltage divider circuit for obtaining logic 1 and logic 2 voltage levels. In incrementor When input S is at logic *0*, Tp_4, Tp_5, and Tp_7 are ON; due to this $V_{dd}/3$ voltage node is connected to the output. When input S is at logic *1*, Tp_5, Tp_6 and Tp_7 are ON; due to this $2V_{dd}/3$ voltage node is connected to the output node. When input S is at logic *2*, Tp_6, Tp_7, and Tp_8 are ON; due to this V_{dd} voltage node is connected to the output node. When input S is at logic 3, Tn_1, Tp_6 and Tp_8 are ON; due to this ground node is connected to the output node.

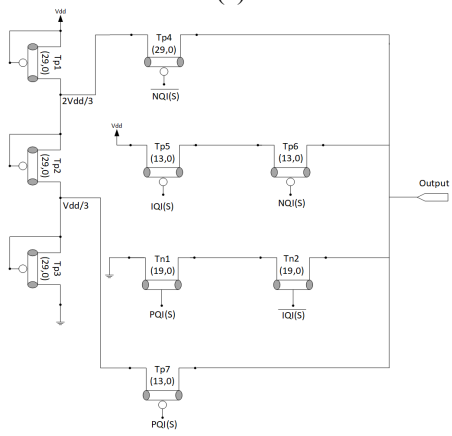
Fig. 3 (b) shows the CNFET implementation of level-II incrementor. When input S is at logic *0*, Tp_4, Tp_5 and Tp_7 are ON while Tp_6, Tn_1 and Tn_2 are OFF; due to this $2V_{dd}/3$ voltage node is connected to the output node. When input S is at logic *1*, Tp_5, Tp_6 and Tn_2 are ON; due to this V_{dd} voltage node is connected to the output node. When input S is at logic *2*, Tp_6, Tn_1, and Tn_2 are ON; due to this ground node is connected to the output node. When input S is at logic *3*, Tp_6, Tp_7, and Tn_1 are ON; due to this $V_{dd}/3$ voltage node is connected to the output node.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:27:03 UTC from IEEE Xplore. Restrictions apply.

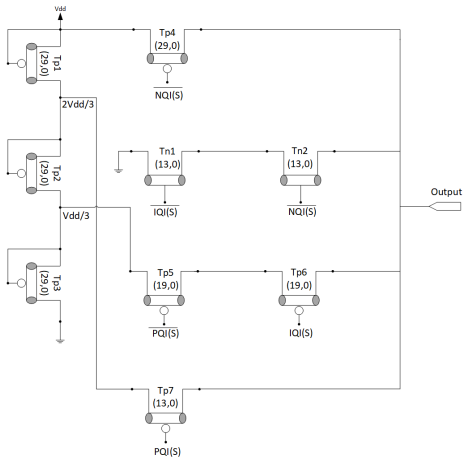
Fig. 3(b) shows the CNFET implementation of decrementor whose working is as follows. When input S is at logic 0, Tp4, Tp5 and Tn1 are ON; due to this Vdd voltage node is connected to the output node. When input S is at logic 1, Tp5, Tn1 and Tn2 are ON; due to this ground node is connected to the output node. When input S is at logic 2, Tp5, Tp6 and Tn2 are ON; due to this Vdd/3 voltage node is connected to the output node. When input S is at logic 3, Tp6, Tp7 and Tn2 are ON; due to this 2Vdd/3 voltage node is connected to the output node.



(a)



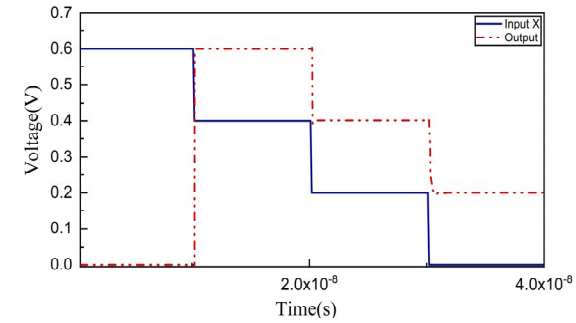
(b)



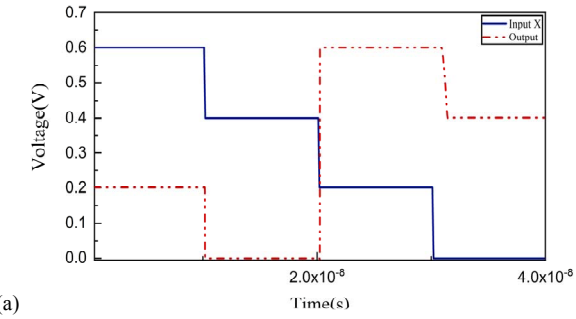
(c)

Fig. 3. CNFET implementation of QTL (a) Incrementor, (b) Level II Incrementor and (b) Decrementor

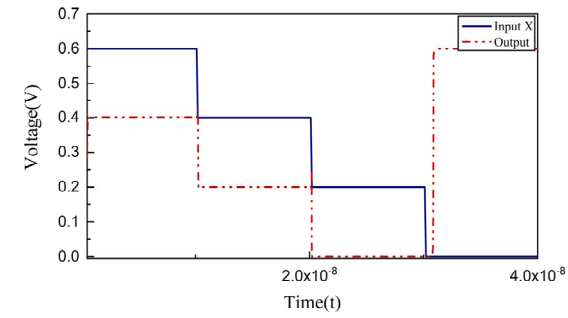
The output waveform response for QTL function incrementor, level-II incrementor and decrementor are given in Fig. 4.



(a)



(b)



(c)

Fig. 4. Output Waveform of QTL (a) Incrementor, (b) Level II Incrementor and (b) Decrementor

B. *Proposed QTL Full Adder*

In this section he novel QTL full adder design is presented or which the truth-table is given in Table III.

Fig. 5 (a) shows the implementation of the proposed QFA sum circuit which contains one 4:1 multiplexer (M_5) with 48 CNFETs, four 2:1 multiplexers (M_1, M_2, M_3 and M_4) with 6 CNFETs each, along with incrementor, Level II incrementor and decrementor consisting of 45 CNFETs each. Therefore, we require 207 CNFETs for implementing the QTL full adder.

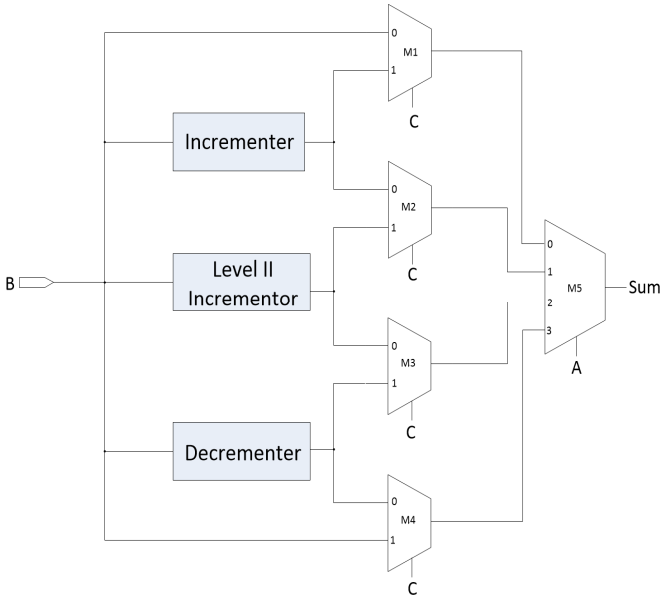
In proposed QTL full adder CNFET implementation, mux M_5 is controlled by input A and multiplexers M_1-M_4 are controlled by input C. Input for multiplexer M_5 is derived from multiplexers M_1-M_4 while for multiplexers M_1-M_4 inputs are derived directly from input B or through the incrementor, level II incrementor or decrementor.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:27:03 UTC from IEEE Xplore. Restrictions apply.

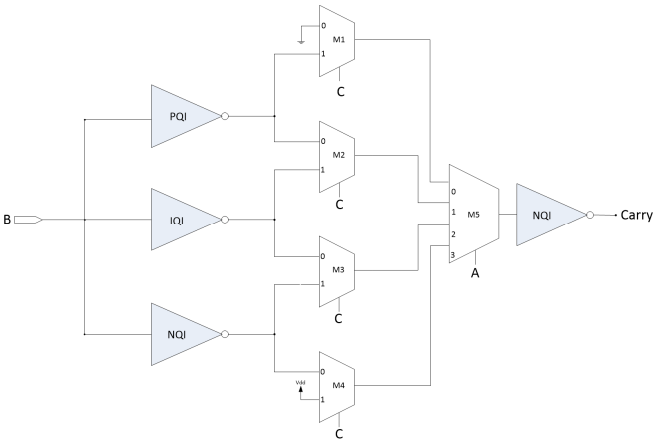
The carry function (C_out) is implemented as given in Fig. 5 (b). For its implementation, we have used 80 CNFETs (one 4:1 multiplexer (M5) with 48 CNFETs, two 2:1 multiplexers (M1, M2, M3 and M4) with 6 CNFETs each, two NQIs with 2 CNFETs each and single PQI and IQI with 2 CNFETs each). Similar to the QFA sum implementation multiplexer M5 is controlled by input A and multiplexers M1-M4 are controlled by input C while, the inputs for multiplexers M1-M4 are derived from the inverter function (PQI, IQI and NQI) of input B or ground or Vdd.

TABLE III. THE TRUTH TABLE FOR QTL ADDER

A	B	Sum		Cout	
		Cin=0	*Cin=1*	*Cin=0*	*Cin=1*
0	0	0	1	0	0
0	1	1	2	0	0
0	2	2	3	0	0
0	3	3	0	0	1
1	0	1	2	0	0
1	1	2	3	0	0
1	2	3	0	0	1
1	3	0	1	1	1
2	0	2	3	0	0
2	1	3	0	0	1
2	2	0	1	1	1
2	3	1	2	1	1
3	0	3	0	0	1
3	1	0	1	1	1
3	2	1	2	1	1
3	3	2	3	1	1



(a)



(b)

Fig. 5. CNFET implementation of QTL (a) Sum generation circuit and (b) Carry generation circuit

Operation explained above can be verified with QFA truth table given in table III with output displayed in Fig. 6. Where In A, In B and In C_in represent input A, input B and input Cin respectively, while output sum and Cout are represented using Sum and Carry. In this 0 V, 0.2 V, 0.4 V and 0.6 V are taken as logic level 0, 1, 2 and 3 respectively.

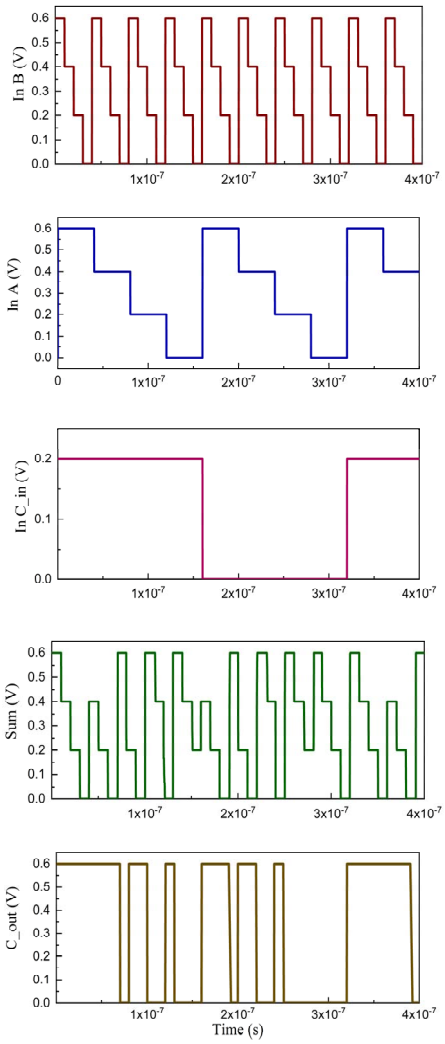


Fig. 6. The output waveform of QTL (a) Sum generation circuit and (b) Carry generation circuit

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:27:03 UTC from IEEE Xplore. Restrictions apply.

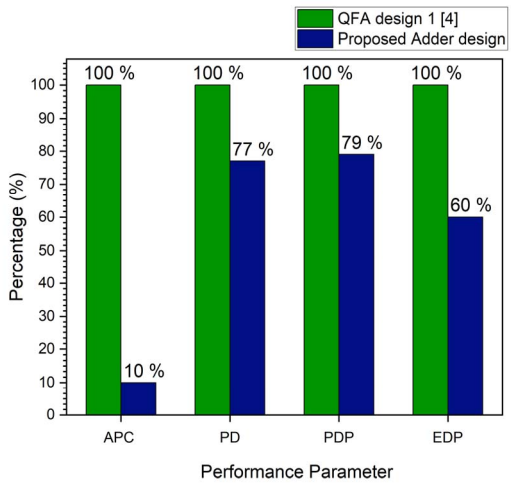
IV. Result and Discussion

From Table IV we can conclude that the proposed QTL full adder design is better than the existing design [4]. Proposed design required total 287 CNFET (207 for sum generation and 80 for carry generation) while in QFA design 1[4] required 356 CNFET. Also QFA design 1[4] shows improvement than the previous works [6]–[8] in terms of PDP and EDP.

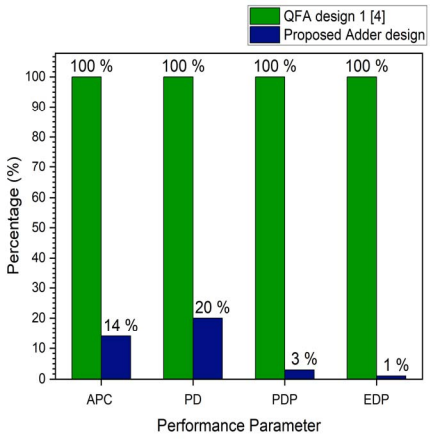
Performance parameters (power, delay, PDP and EDP) for the proposed QTL full adder design are calculated for two types of load conditions i.e. *C_load* = 1 *fF* and *C_load* = 2 *fF*. The proposed design is ~90% better in terms of power and ~22% (for *C_load*= 1 *fF*) & ~80% (for *C_load*= 2 *fF*) better in terms of propagation delay than the existing QFA design 1 [3]. Similarly, in the case of PDP and EDP, the proposed design is approximately 99% better than the existing design. In QFA design 1 three power supply (Vdd, Vdd/2 and Vdd/3) are used while in proposed design only 1 power supply (Vdd) is needed. Fig. 7 (a) and (b) depict the comparison between the performance parameters of the proposed QTL full adder with QFA design 1 [4] for *C_load* 1 *fF* and 2 *fF* (in %).

TABLE IV. Performance Analysis for QTL Full Adder.

Parameter	C_load	Proposed Adder design	QFA design 1 [4]
Average Power Consumption (APC) (uW)	1e-15f	0.0217 (10%)	0.212 (100%)
	2e-15f	0.0589 (14%)	0.453 (100%)
Propagation delay (PD) (ps)	1e-15f	8.458 (77%)	10.92 (100%)
	2e-15f	54.626 (20%)	277.6 (100%)
PDP (e-16 J)	1e-15f	0.001801 (79%)	0.00231 (100%)
	2e-15f	0.03217 (3%)	1.2575 (100%)
EDP (e-28Js)	1e-15f	0.01523286 (60%)	0.02523 (100%)
	2e-15f	1.7573184 (1%)	349.082 (100%)
Power supply Count		1	3



(a)



(b)

Fig. 7. Comparison of performance parameters for Proposed QTL Adder with QFA design 1 [3] for C-load (a) 1fF and (b) 2fF (in %)

V. Conclusion

In this work, novel QFA is implemented using incrementor, level II incrementor, and decrementor based on CNFET.

Performance parameters calculated for C-load 1fF & 2fF indicate that the proposed design of QTL full adder is approximately *90%* and *94%* efficient respectively in terms of APC and *22%* (for *Cload= 1fF*) & *~80%* (for *C_load= 2fF*) efficient in terms of PD than the previous designs. PDP value for the proposed adder design is approximately *21%* (for *C_load= 1fF*) and *97%* (for *C_load= 2fF*) less than the previous design. Also, the EDP of the proposed full adder is *40%* (for *C_load= 1fF*) and *99%* (for *C_load= 2fF*) better. Moreover, the number of power supplies required is reduced from *3* to *1*. In the Proposed QFA design CNFET count decreased from 356 to 287. Evaluation of the performance parameters under different load conditions suggests that the proposed designs are more efficient to implement the QFA than existing designs.

References

[1] J. A. Carballo, W. T. J. Chan, P. A. Gargini, A. B. Kahng, and S. Nath, "ITRS 2.0: Toward a re-framing of the Semiconductor Technology Roadmap," 2014 32nd IEEE Int. Conf. Comput. Des. ICCD 2014, pp. 139–146, 2014.

[2] Z. Chen, L. Wei, A. Keshavarzi, and K. Roy, "I DDQ testing for deep-submicron ICs: Challenges and solutions," IEEE Des. Test Comput., vol. 19, no. 2, pp. 24–33, 2002.

[3] D. E. Nikonov and I. A. Young, "Overview of beyond-CMOS devices and a uniform methodology for their benchmarking," Proc. IEEE, vol. 101, no. 12, pp. 2498–2533, 2013.

[4] E. Roosta and S. A. Hosseini, "A Novel Multiplexer-Based Quaternary Full Adder in Nanoelectronics," Circuits, Syst. Signal Process., vol. 38, no. 9, pp. 4056–4078, 2019.

[5] M. Kameyama, "TOWARD THE AGE OF ' BEYOND-BINARY ELECTRONICS Multiple- Valued Integrated Circuits in Hardware Algorithms Based Representation," no. 1, pp. 1–5.

[6] K. S. Vasundara Patel and K. S. Gurumurthy, "Design of high performance quaternary adders," Proc. - 41st IEEE Int. Symp. Mult. Logic, ISMVL 2011, pp. 22–26, 2011.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:27:03 UTC from IEEE Xplore. Restrictions apply.

[7] S. A. Ebrahimi, M. R. Reshadinezhad, A. Bohlooli, and M. Shahsavari, "Efficient CNTFET-based design of quaternary logic gates and arithmetic circuits," Microelectronics J., vol. 53, pp. 156–166, 2016.

[8] F. Sharifi, M. H. Moaiyeri, K. Navi, and N. Bagherzadeh, "Quaternary full adder cells based on carbon nanotube FETs," J. Comput. Electron., vol. 14, no. 3, pp. 762–772, 2015.

[9] F. Sharifi, M. H. Moaiyeri, H. Sharifi, K. Navi, and H. Thapliyal, "On the design of quaternary arithmetic logic unit based on CNTFETs," Int. J. Electron. Lett., vol. 7, no. 1, pp. 1–13, 2019.

[10] J. Deng and H.-. P. Wong, "A Compact SPICE Model for Carbon-Nanotube Field-Effect Transistors Including Nonidealities and Its Application—Part I: Model of the Intrinsic Channel Region," IEEE Trans. Electron Devices, vol. 54, no. 12, pp. 3186–3194, Dec. 2007.

[11] D. Jie and H. S. P. Wong, "A Compact SPICE Model for Carbon-Nanotube Field-Effect Transistors Including Nonidealities and Its Application—Part II: Full Device Model and Circuit Performance Benchmarking," Electron Devices, IEEE Trans., vol. 54, no. 12, pp. 3195–3205, 2007.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:27:03 UTC from IEEE Xplore. Restrictions apply.

arXiv:2009.01571v1 [cs.LG] 3 Sep 2020

MixBoost: Synthetic Oversampling with Boosted Mixup for Handling Extreme Imbalance

Anubha Kabra*§, Ayush Chopra‡, Nikaash Puri†, Pinkesh Badjatiya†, Sukriti Verma†, Piyush Gupta† and Balaji K†
†Media and Data Science Research Lab, Adobe
‡Massachusetts Institute of Technology
§Delhi Technological University

***Abstract*—Training a classification model on a dataset where the instances of one class outnumber those of the other class is a challenging problem. Such imbalanced datasets are standard in real-world situations such as fraud detection, medical diagnosis, and computational advertising. We propose an iterative data augmentation method, *MixBoost*, which intelligently selects ($Boost$) and then combines (Mix) instances from the majority and minority classes to generate synthetic hybrid instances that have characteristics of both classes. We evaluate *MixBoost* on 20 benchmark datasets, show that it outperforms existing approaches and test its efficacy through significance testing. We also present ablation studies to analyze the impact of the different components of MixBoost.**

***Index Terms*—imbalanced datasets, neural networks, sampling, minority sampling, data augmentation, class imbalance**

I. Introduction

Several real-world situations (fault detection [23], disease classification [16], software failures [4], oil spill detection [15] and protein sequence detection [2]) involve learning from datasets where the instances of one class (the majority class) far outnumber those of the other class (the minority class). In certain **extremely imbalanced** datasets, the ratio of the number of instances from the two classes is very low ($< 1:100$) or the number of instances of the minority class is very small [18]. For instance, in gamma-ray anomaly detection [26], the datasets for learning have 10 anomalies and approximately 25,000 benign signatures. In fraud detection, there are typically about 5 fraudulent examples in a dataset of over 300,000 transactions [28].

Training a binary classification model on such extremely imbalanced datasets is a challenging problem. A popular class of methods alleviates this problem by augmenting the training data with synthetic instances before training the classification model [5], [6], [12], [18]. These data augmentation methods generate synthetic minority class instances using either minority or majority class instances from the training dataset. However, existing methods often generate instances that don't improve (or even worsen) the classifier performance. Intuitively, these methods generate good quality synthetic instances in regions of the input space where the classification model is already accurate. In regions where the model is inaccurate and requires synthetic instances, these methods often do not perform as well. Figure 1 illustrates this intuition. We believe that using the trained model (whose performance we are trying to improve) to guide the data augmentation process will allow us to augment regions of the input space where the model performs poorly. Further, existing methods generate synthetic **homogeneous** instances, i.e., instances that belong to a single class (usually the minority class). Recent work in Computer Vision [19], [21], [24] has demonstrated the value of augmenting training datasets with non-homogeneous **hybrid** instances to learn more robust representations.

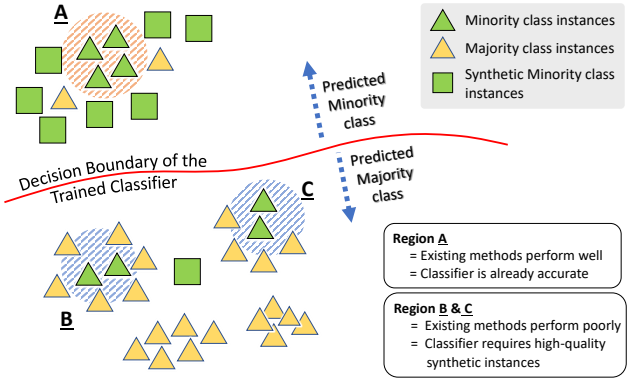


Fig. 1: Illustration to describe the limitation of synthetic oversampling techniques. Existing methods (SMOTE [5] and its variants, SWIM [18]) select candidate instances from one of the classes and create synthetic instances based on these selected instances. Therefore, most generated synthetic instances lie near clusters (often within the convex hull) of instances that are often already correctly classified by the classification model (region A). Further, these methods generate fewer and poorer quality synthetic instances in regions of the input space where the model does not perform well (regions B and C).

We describe a data augmentation method (*MixBoost*) that improves classifier performance on such imbalanced datasets. Our key contributions are:

- We introduce a data augmentation method, *MixBoost* that generates synthetic **hybrid** instances to augment an imbalanced dataset (Section II). *MixBoost* has two innovative components. **First**, it mixes instances of the minority and majority classes to generate synthetic hybrid

*Work done during internship at MDSR
‡Work done while at Adobe

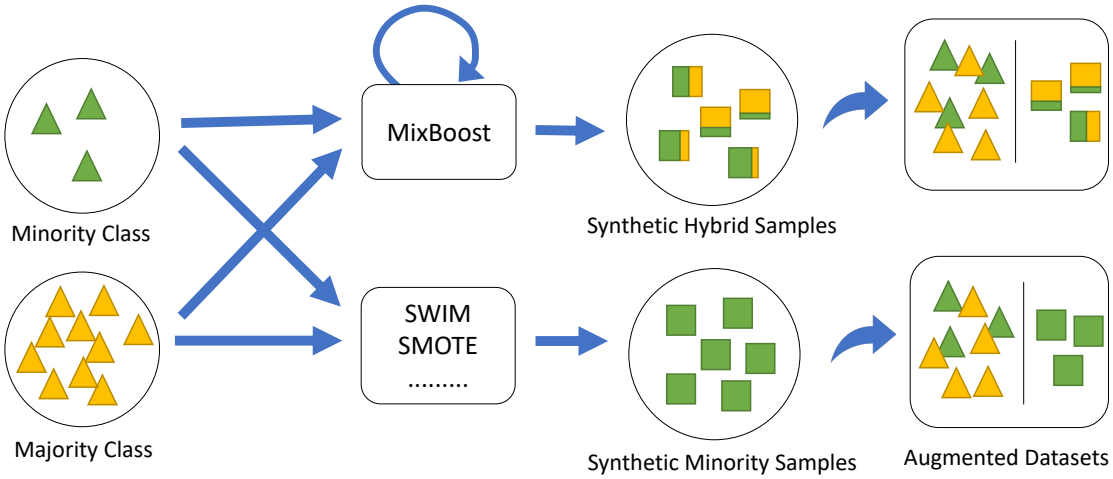


Fig. 2: High-level summary of the data-augmentation pipeline. The triangle represents the data instances in the non-augmented training dataset. Green is for minority class and yellow for majority class. The square represents the synthetic generated instances that are added to the training dataset (augmentation) prior to training. Traditional approaches such as SWIM [18], SMOTE [5], and its variants generate synthetic minority class instances (denoted by green). In contrast, *MixBoost* generates synthetic hybrid instances that interpolate the minority and majority class instances (denoted by the green-yellow color scheme).

instances that have elements of both classes (Section II-D). **Second**, it intelligently selects the instances for mixing using a novel entropy-weighted **low-high** technique (Section II-C).

- We show that *MixBoost* outperforms state-of-the-art data augmentation methods on multiple highly imbalanced benchmark datasets for different levels of class imbalance (Section IV).

There are several categories of methods to augment an imbalanced dataset prior to training a classification model. Under-sampling methods discard instances of the majority class at random to balance the class distribution. However, removing instances of the majority class can lead to a loss of information and subsequently degrade classifier performance. Over-sampling methods duplicate instances of the minority class at random to balance the class distribution. These methods add duplicates of existing instances to the training dataset and therefore do not add new information [10], [11]. More sophisticated data augmentation methods augment the training datasets by creating synthetic minority class instances. For instance, SMOTE [5] creates synthetic minority instances by interpolating minority class instances in the training data. However, SMOTE ignores majority class data when creating synthetic instances. Variants of SMOTE (Borderline SMOTE [12], ADASYN [14], SMOTEBoost [6]) use the distribution of majority class data to filter instances created using minority class data. However, since these methods use only minority class data, the synthetic instances they create are restricted to the convex-hull of the minority class distribution. To expand the space spanned by generated instances, SWIM [18] creates instances by inflating minority class data along the density contours of majority class data.

Existing data augmentation methods give promising results in a variety of situations. However, these methods focus on either the majority class or the minority class. Approaches that focus on the minority class often overlook the majority class and can degrade classifier performance by generating borderline or overlapping instances. On the other hand, approaches that focus on the majority class tend to overfit the training distribution. In contrast, *MixBoost* combines instances from the majority and minority classes to create synthetic hybrid instances. Each synthetic hybrid instance contains elements of a majority class and a minority class instance. Further, *MixBoost* uses an information-theoretic measure along with the trained classifier to select the instances to combine. Figure 2 illustrates the essential idea and distinction to related approaches.

Figure 3 shows the high-level architecture of *MixBoost*. *MixBoost* generates synthetic hybrid instances iteratively. Each iteration consists of a $Boost$ step followed by a Mix step. In the $Boost$ step, *MixBoost* uses the trained classifier to intelligently select pairs of instances from the minority and majority class for mixing. Intuitively, *MixBoost* selects pairs of instances such that one instance in each pair is close to the decision boundary of the trained classifier, and the other instance is far from the decision boundary. In the Mix step, *MixBoost* mixes the selected instances to create synthetic hybrid instances that are used to augment the training dataset. The classifier is re-trained with the original data augmented with the hybrid instances before the next iteration of *MixBoost*.

In Section II we delineate the proposed approach. Next, in Section III we present our experimental setup and evaluation strategy. In Section IV we compare quantitative performance of *MixBoost* with multiple strong baselines on 20 benchmark datasets as well as test the efficacy of our framework through significance tests. In Section V we perform ablation studies to analyze the impact of different design choices for the proposed

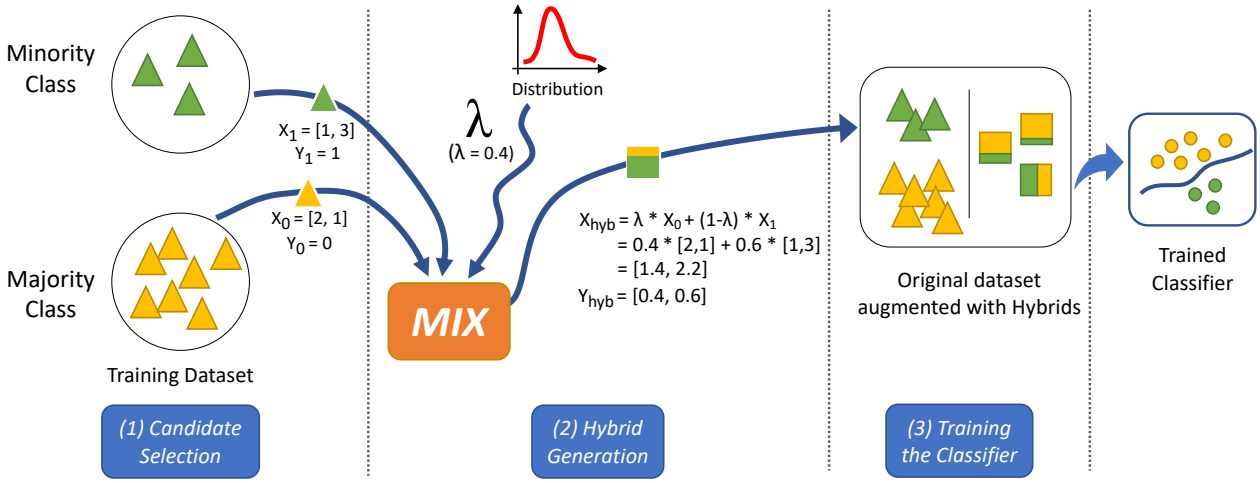


Fig. 3: High-level architecture of *MixBoost*. In the (1) candidate selection step ($Boost$), *MixBoost* uses entropy to intelligently select instances from the training dataset. In the (2) hybrid generation step (Mix), we mix the selected instances to create synthetic hybrid instances. λ is the interpolation ratio that controls the extent of overlap between the instances to be mixed. We then augment the training dataset with the generated hybrid instances and train a classifier on the augmented dataset.

approach. Finally, we summarize related work in Section VI

II. Approach

In this section, we explain in detail the various steps of our proposed *MixBoost* algorithm. In Section II-A we define the terminologies used throughout the paper. In Section II-B we provide a high-level overview of the proposed approach followed by detailed explanation of the different stages in Sections II-C and II-D.

A. Definitions

A binary classification task is characterized by a dataset $X^{n \times m} \in R$ and corresponding labels $Y^n \in \{0, 1\}$. The class label 0 corresponds to the majority class, and label 1 corresponds to the minority class throughout the paper. Let n be the number of instances in the dataset, and m be the number of features for an instance in the dataset. Then, the training objective is to learn a function, $f(x^i) \rightarrow y^i$, that maps the instances $x^i \in X$ to their corresponding labels $y^i \in Y$. For imbalanced datasets, the number of instances of the majority class far outnumbers that of the minority class. The dataset, denoted by $D^{orig} := (X, Y)$, is split into a training dataset $D^{orig}_{train} := (X_{train}, Y_{train})$, and a testing dataset $D^{orig}_{test} := (X_{test}, Y_{test})$. Individual datapoints from these datasets are referred to as d^{orig}_{train} and d^{orig}_{test} respectively. Let n_{train} be the number of instances in the training dataset. We train a binary classification model M on D^{orig}_{train} and evaluate on the test dataset D^{orig}_{test} using $g\text{-}mean$ [15] and ROC-AUC [25] metrics.

MixBoost uses the classifier M along with dataset D^{orig}_{train} to create synthetic hybrid instances. The instances are added to the training dataset to create the augmented dataset D^{aug}_{train} which is then used to in-turn train the classifier M. This represents a single iteration of *MixBoost*. *MixBoost* is run for k such iterations to expand the augmented dataset. We report performance by evaluating the model trained using the final augmented dataset, obtained at the end of k iterations, using D^{orig}_{test}.

B. MixBoost: High-Level Summary

MixBoost generates synthetic hybrid instances by interpolating instances from the majority and minority classes. A single iteration of *MixBoost* (summarized in Figure 3) consists of the following steps:

1) **Candidate Selection ($Boost$):** We sample candidate instances from the majority and minority classes in D^{orig}_{train} prior to mixing. We experiment with both random and guided sampling techniques (Section II-C).
2) **Hybrid Generation (Mix):** We mix the sampled instances to generate synthetic hybrid instances (Section II-D). $D^{hyb} := (X^{hyb}, Y^{hyb})$ denotes the set of synthetic instances generated in this step (and each synthetic instance is correspondingly referred to as d^{hyb}).

At the end of a *MixBoost* iteration, we add D^{hyb} to D^{aug}_{train} and retrain the classifier M on D^{aug}_{train}.

For ease of explanation, we focus on generating a single synthetic hybrid instance. However, the steps to generate several instances follow from this single instance case. *MixBoost* creates a synthetic hybrid instance by mixing a majority class instance (x_0) and a minority class instance (x_1). There are *two* important considerations to this process. First, how should *MixBoost* select the *best* candidates x_0 and x_1 from D_{train}? Second, how should *MixBoost intelligently* combine x_0 and x_1 to create the synthetic hybrid instance d^{hyb}? We first consider the problem of instance selection.

C. MixBoost: Candidate Selection (the $Boost$ step)

Unlike existing works, *MixBoost* creates synthetic datapoints that are conditioned on both the majority and minority class instances. For generating these synthetic hybrids, we propose two alternative strategies for selection of the candidate instances,

1) ***R-Selection****:* First, we propose *R-Selection* which randomly selects (majority and minority) candidates x_0 and x_1 from D^{orig}_{train} using uniform random variables. The probability of selecting an instance is equal for all instances within the majority and minority classes. Formally, if p^i_0 is the probability of selecting the ith majority class instance and p^j_1 is the probability of selecting the jth minority class instance, then:

$$p^i_0 = 1/n_0 \ \forall i, \qquad p^j_1 = 1/n_1 \ \forall j \tag{1}$$

where n_0 is the number of majority class instances and n_1 is the number of minority class instances in D^{orig}_{train}.

2) ***Entropy Weighted (EW) Selection****: R-Selection* assumes that all synthetic hybrid instances generated by mixing any combination of minority/majority class instances are equally helpful additions to the augmented dataset for training. However, it is possible that mixing certain combinations of minority/majority class instances from D^{orig}_{train} can result in better training performance.

For a given classification model M, entropy is a measure of uncertainty in the models prediction. Consequently, we introduce an alternate strategy called Entropy-Weighted selection (*EW-Selection*). *EW-Selection* posits that selecting candidates from D^{orig}_{train} by actively weighting them with the uncertainty in their model predictions will result in generation of more useful hybrid instances. Formally, if $\vec{y}^{\ i}_{pred}$ is the probability distribution over target classes output by the classifier M for the ith instance in D^{orig}_{train}, then the entropy of the sample, denoted by E^i, is computed as,

$$E^i = Entropy(\vec{y}^{\ i}_{pred}) = -\sum y^i_{pred} * log(y^i_{pred}) \tag{2}$$

where the summation is over the elements of the vector $\vec{y}^{\ i}_{pred}$.

For a given data instance, a high E^i implies that model M is less certain about the ground-truth class prediction of the instance while a low E^i implies that M is more certain about the ground-truth class prediction of the instance. We hypothesise that augmenting training dataset with synthetic instances in vicinity of high entropy feature sub-spaces (where model is currently uncertain about predictions) can improve model training performance.

Let $\hat{E_0} = \sum E^i \ \forall x^i_0$ (x^i with class c_0) is the sum of entropy values for majority class instances and $\hat{E_1} = \sum E^i \ \forall x^i_1$ (x^i with class c_1) denote the sum of entropy values for minority class instances. Then, $P(x^i|c_0)$ and $P(x^j|c_1)$ denote the *entropy ratios*, which are the probability of selecting a majority class instance and a minority class instance respectively for mixing. Both $P(x^i|c_0)$ and $P(x^j|c_1)$ are computed as,

$$P(x^i|c_0) = \frac{E^i}{\hat{E_0}}, \qquad P(x^j|c_1) = \frac{E^j}{\hat{E_1}} \tag{3}$$

```python
from over_sampling import MixBoost
import numpy


X_train, Y_train = get_date()     #returns training dataset

interp_dist = lambda: numpy.random.beta(0.5, 0.5)

mb = MixBoost(interp_dist=interp_dist, num_iters=10,\
                    random_state=42)

X_aug, Y_aug = mb.fit_resample(X_train, Y_train)
```

Fig. 4: Code snippet that illustrates how to use $MixBoost$ in Python.

In practice, we find that selecting one candidate instance with *high entropy ratio* and the other with *low entropy ratio* leads to generation of synthetic hybrid instances that result in the best performance. *MixBoost* interleaves this **low-high** selection with the generation of hybrid instances using the Mix step described below.

D. MixBoost: Hybrid Generation (the Mix step)

Let (x_0, y_0) and (x_1, y_1) be instances selected (in the $Boost$ step) from the majority and minority class respectively. We adopt the mixing strategy introduced in [24].

$$\begin{aligned} x_{hyb} &= \lambda * x_0 + (1 - \lambda) * x_1 \\ y_{hyb} &= \lambda * y_0 + (1 - \lambda) * y_1 \end{aligned} \tag{4}$$

(x_{hyb}, y_{hyb}) is the generated synthetic hybrid instance. λ is the interpolation ratio. Intuitively, λ controls the extent of overlap between the two instances used to generate the synthetic hybrid instance. MIXUP [24] and BC+ [21] are existing image augmentation techniques which use linear and non-linear interpolations respectively. BC+ samples λ using the Uniform Distribution $U(0, 1)$ while MIXUP samples λ using the $Beta(\alpha, \alpha)$ distribution. For most tasks, MIXUP uses $\alpha = 1$, which also reduces to sampling λ from a Uniform Distribution [19]. In contrast, *MixBoost* works with skewed data distributions arising from extreme class imbalance. Correspondingly, we explore sampling λ from several probability density functions and discuss our findings in Section V-B.

The resultant synthetic hybrid instances (D^{hyb}) are used to augment the training dataset (D^{orig}_{train}), generating D^{aug}_{train}, which is used to (re)train the classifier. The classifier is (re)trained on a *cross-entropy* loss by using the soft labels for the D^{hyb} and the one-hot labels for D^{orig}_{train} as ground truths. In the subsequent iterations, this classifier is used to update entropy values for candidates in the original training dataset to facilitate the next iteration of *MixBoost*. The entire workflow of our approach is shown in Figure 3.

III. Experimental Settings

A. Datasets

To evaluate *MixBoost* against the state-of-the-art oversampling methods, we compare performance on binary classifi-

cation on **20 benchmark datasets** (Table I, available here[1]). To ensure the evaluation consistency we use the same dataset and imbalance ratios as in [18]. For robust evaluation, we split the dataset into equal training and testing halves. We randomly down-sample the minority class instances in the training dataset to simulate different levels of extreme imbalance. Specifically, we test at three levels of imbalance, with minority training dataset sizes of 4, 7, and 10 to simulate the extreme imbalance scenarios [18]. We further skew the data distribution (2 & 3 minority class instances) to show that *MixBoost* outperforms existing methods when using even fewer number of minority class training instances.

B. Classification

For all experiments, we use the Naive Bayes, Nearest Neighbour, Decision Tree, Support Vector Machine, and Multi-Layer Perceptron (MLP) binary classifiers. For each dataset, we compare the best performance of each data augmentation method across the various binary classifiers against the performance of *MixBoost*. We use the *sklearn* [17] package in Python for our experiments; all classifiers use the default parameters. We use *tensorflow* to implement the MLP. For the MLP, we use *Adam* optimizer, a learning rate of $1e-4$, a batch size of 500, and train it for 300 epochs. We use the same hyper-parameters for all datasets. We normalize the data before training and testing.

C. Data Augmentation

We use each data augmentation method to generate n synthetic instances, where n is the number of instances in the training dataset. For $MixBoost$, we find that generating the n instances over 5 iterations ($0.2n$ in each iteration) leads to the best results. We sample the interpolation ratio λ from a $Beta(0.5, 0.5)$ distribution. We explore the sensitivity of $MixBoost$ to these hyper-parameters in Section V.

D. Evaluation

We compare data augmentation methods on two metrics. First, we use $g\text{-}mean$ [15], which is the geometric mean of the True Positive Rate (TPR) (for majority class) and the True Negative Rate (TNR) (for minority class).

$$g\text{-}mean = \sqrt{TPR \times TNR} \tag{5}$$

We use $g\text{-}mean$ because it is both immune to imbalance [15] and provides information about extreme imbalance [18]. Second, for completeness, we also compare methods using the ROC-AUC scores. We repeat all experiments 30 times and report the mean and standard deviation (rounded to the nearest decimal place) of the metrics for each dataset. Figure 4 includes a code snippet illustrating *MixBoost* usage.

We compare *MixBoost* to Random Over-sampling and Under-sampling (ROS, RUS), SMOTE [5], Borderline-SMOTE (B1, B2) [12], SMOTE with Tomek Links [15], ADASYN [14] and SWIM [18]. To make reporting easier, we merge all re-sampling approaches (ROS, RUS, SMOTE, B1, B2, SMOTE with Tomek Links, ADASYN) as Alternative Re-sampling Techniques (ALT) and report the best performing combination of data augmentation method and classifier for each dataset. We report the results from SWIM [18] separately since it is the most recent work, and it outperforms methods in ALT on most datasets. Finally, $Baseline$ represents the case where no data augmentation is used before training the binary classifier.

[1]Datasets can be downloaded from [7]

Dataset	Features	No. of majority instances	R4	R7	R10
Abalone 9-18	8	689	1:173	1:99	1:69
Diabetes	8	500	1:125	1:72	1:50
Wisconsin	9	444	1:111	1:64	1:456
Wine Q. Red 4	11	1546	1:387	1:221	1:155
Wine Q. White	11	880	1:220	1:126	1:88
Vowel 10	13	898	1:225	1:129	1:90
Pima Indians	8	500	1:125	1:72	1:50
Vehicle 0	18	641	1:160	1:91	1:64
Vehicle 1	18	624	1:156	1:89	1:62
Vehicle 2	18	622	1:155	1:88	1:62
Vehicle 3	18	627	1:156	1:89	1:62
Ring Norm	20	3736	1:934	1:534	1:374
Waveform	21	600	1:150	1:86	1:60
PC4	37	1280	1:320	1:183	1:128
Piechart	37	644	1:161	1:92	1:65
Pizza Cutter	37	609	1:153	1:87	1:61
Ada Agnostic	48	3430	1:858	1:490	1:343
Forest Cover	54	2970	1:743	1:425	1:297
Spam Base	57	2788	1:697	1:399	1:279
Mfeat Karhu.	64	1800	1:450	1:258	1:180

TABLE I: Description of the datasets used in our experiments. To ensure evaluation consistency, we use the same datasets and configuration as proposed by [18]. R4, R7, and R10 denote the ratio of class imbalance (minority:majority) after down-sampling the training datasets to have 4, 7, and 10 minority class instances respectively to simulate the extreme imbalance [18] scenarios (as discussed in Section I)

IV. Results

In this section, we present quantitative analysis of the proposed *MixBoost* oversampling technique. First, we compare *MixBoost* with existing state-of-the-art techniques and report results using the $g\text{-}mean$ metric (Table II). Next, for completeness of study, we report comparative performance on the ROC-AUC metric (Table IV), contrast the different *MixBoost* instance selection strategies (Table III) and perform significance tests for different levels of imbalance (Section IV-B).

A. Quantitative Results

As in previous works, we compare *MixBoost* with existing state-of-the-art techniques in Table II using the $g\text{-}mean$ metric, which is considered a better metric for evaluating performance when dealing with extreme imbalance in datasets [15], [18]. The results shown here are obtained by down-sampling to the most extreme imbalance setting R4 where there are only 4 minority class instances (see Table I). As evident, *MixBoost* outperforms the existing methods on **17 out of the**

Dataset	Baseline	ALT	SWIM	*MixBoost*
Abalone 9-18	0.481	0.612	0.723	**0.743**
Diabetes	0.259	0.509	0.509	**0.701**
Wisconsin	0.874	0.956	0.958	**0.969**
Wine Q. Red 4	0.224	0.502	0.535	**0.815**
Wine Q.White 3v7	0.451	0.572	0.730	**0.750**
Vowel 10	0.724	0.738	0.812	**0.845**
Pima Indians	0.276	0.479	0.509	**0.700**
Vehicle 0	0.534	0.758	0.814	**0.900**
Vehicle 1	0.541	0.739	**0.791**	0.735
Vehicle 2	0.450	0.549	0.560	**0.880**
Vehicle 3	0.402	0.505	0.569	**0.651**
Ring Norm	0.274	**0.933**	0.899	0.580
Waveform	0.301	0.701	0.688	**0.844**
PC4	0.572	0.559	0.611	**0.737**
PieChart	0.455	0.516	0.576	**0.741**
Pizza Cutter	0.468	0.506	0.552	**0.725**
Ada Agnostic	0.451	0.445	0.539	**0.690**
Forest Cover	0.561	0.554	0.550	**0.917**
Spam Base	0.440	0.550	0.685	**0.872**
Mfeat Karhunen	0.274	**0.933**	0.899	0.927

TABLE II: Comparative $g\text{-}mean$ results (mean) for *MixBoost* with existing over-sampling methods. These results represent the R4 setting where the training dataset has 4 minority class instances. Baseline refers to the case where classifier is trained without data augmentation. ALT is the score of the best performing data augmentation strategy (other than SWIM [18] and *MixBoost*) as described in Section III-D. The best score for each dataset is highlighted in **bold**.

Dataset	*MixBoost*	
	R-Selection	**EW-Selection**
Abalone 9-18	0.743 ± 0.03	0.735 ± 0.06
Diabetes	0.701 ± 0.05	0.560 ± 0.04
Wisconsin	0.960 ± 0.02	0.969 ± 0.08
Wine Q. Red 4	0.714 ± 0.04	0.815 ± 0.08
Wine Q. White 3v7	0.743 ± 0.06	0.750 ± 0.05
Vowel 10	0.845 ± 0.04	0.854 ± 0.07
Pima Indians	0.700 ± 0.05	0.597 ± 0.04
Vehicle 0	0.900 ± 0.05	0.850 ± 0.02
Vehicle 1	0.700 ± 0.03	0.735 ± 0.03
Vehicle 2	0.880 ± 0.02	0.638 ± 0.03
Vehicle 3	0.651 ± 0.06	0.600 ± 0.03
Ring Norm	0.550 ± 0.04	0.580 ± 0.03
Waveform	0.812 ± 0.03	0.844 ± 0.05
PC4	0.720 ± 0.08	0.737 ± 0.04
PieChart	0.611 ± 0.06	0.741 ± 0.07
Pizza Cutter	0.725 ± 0.05	0.678 ± 0.07
Ada Agnostic	0.690 ± 0.02	0.648 ± 0.03
Forest Cover	0.910 ± 0.05	0.917 ± 0.02
Spam Base	0.872 ± 0.03	0.834 ± 0.05
Mfeat Karhunen	0.888 ± 0.07	0.927 ± 0.05

TABLE III: Comparative $g\text{-}mean$ results (mean and standard deviation for 30 independent runs) for the different candidate selection strategies (refer to Section IV for details) of our approach *MixBoost*.

20 datasets. SWIM achieves best performance on only one while ALT achieves superior performance on 2 datasets. For each dataset, ALT reports the best performance amongst all algorithms other than SWIM and *MixBoost* which are stated in III-D. In *MixBoost*, the best results of the *R-Selection* and *EW-Selection* strategies are reported on each dataset. Table III compares the performance of the different selection strategies used in *MixBoost*. The relative performance of *EW-Selection* and *R-Selection* strategies for *MixBoost* indicates that leveraging an entropy prior during the candidate selection leads to the creation of more useful synthetic hybrid candidates than selecting instances at random, in several cases. Interestingly, both *R-Selection* and *EW-Selection* strategies independently outperform existing methods on **17 out of 20 datasets**.

Table IV reports comparative results with the ROC-AUC metric. For ease of analysis, we restrict the comparison to SWIM since it is the strongest baseline which consistently outperforms other techniques and is also the most recent state-of-the-art. The results highlight that *MixBoost* outperforms SWIM on the ROC-AUC metric as well by achieving superior performance on **20 out of 20 datasets**.

To study the sensitivity of *MixBoost*, we analyse and compare standard deviation for different combinations of *MixBoost* along with SWIM in Tables III and IV. The results show that *MixBoost* results in a mean standard deviation (over all 20

Dataset	SWIM	*MixBoost*
Abalone 9-18	0.790 ± 0.08	**0.820** ± 0.06
Diabetes	0.778 ± 0.03	**0.800** ± 0.02
Wisconsin	0.899 ± 0.08	**0.972** ± 0.01
Wine Q. Red 4	0.950 ± 0.03	**0.971** ± 0.03
Wine Q. White 3 vs 7	0.640 ± 0.09	**0.783** ± 0.05
Vowel 10	0.895 ± 0.05	**0.958** ± 0.03
Pima Indians	0.778 ± 0.03	**0.800** ± 0.02
Vehicle 0	0.628 ± 0.04	**0.896** ± 0.07
Vehicle 1	0.592 ± 0.06	**0.750** ± 0.02
Vehicle 2	0.824 ± 0.03	**0.921** ± 0.03
Vehicle 3	0.585 ± 0.05	**0.664** ± 0.03
Ring Norm	0.704 ± 0.24	**0.892** ± 0.08
Waveform	0.828 ± 0.02	**0.869** ± 0.02
PC4	0.686 ± 0.09	**0.794** ± 0.03
PieChart	0.661 ± 0.06	**0.743** ± 0.08
Pizza Cutter	0.662 ± 0.08	**0.735** ± 0.04
Ada Agnostic	0.671 ± 0.05	**0.798** ± 0.02
Forest Cover	0.881 ± 0.07	**0.970** ± 0.02
Spam Base	0.769 ± 0.12	**0.835** ± 0.07
Mfeat Karhunen	0.980 ± 0.01	**0.992** ± 0.00

TABLE IV: ROC-AUC scores (mean and standard deviation) for EW-Selection strategy of *MixBoost* and SWIM. The best score for each dataset is highlighted in **bold**.

datasets) of 0.047 which is comparable with 0.044 by SWIM. In the next section, we also probe the statistical significance of our observed results, with respect to the strongest baseline (SWIM), across varying levels of class imbalance.

B. Statistical Significance

We use the Bayesian signed test [3] to evaluate the results presented in the previous section. The Bayesian signed test is used to compare two classification methods over several datasets. We use the test to compare *MixBoost* to SWIM over all 20 datasets. We compare the methods for training datasets down-sampled to 4, 7, and 10 minority class instances. Using the Bayesian method enables us to ask questions about posterior probabilities, that we cannot answer using null hypothesis tests [18].

Concretely, we wish to ascertain, based on the experiments, what is the probability that *MixBoost* is better than SWIM for data augmentation. We repeat the setup described in [18], where, based on the assumption of the Dirichlet process, the posterior probability for the Bayesian signed test is calculated as a mixture of Dirac deltas centered on the observation.

Figure 5 shows the posterior plots of the Bayesian signed test for the three down-sampled datasets. The posteriors are calculated with the prior parameter of the Dirichlet as $s = 0.5$ and $z_0 = 0$ as suggested by [18]. The posterior plots report the samples from the posterior (blue cloud of points), the simplex (the large triangle), and three regions. The three regions of the triangle denote the following, (1) The region in the bottom left of the triangle indicates the case where it is more probable that SWIM is better than *MixBoost*. (2) The region in the bottom right of the triangle indicates the opposite, i.e., when it is more probable that *MixBoost* is better than SWIM. (3) The region at the top of the triangle indicates that it is likely that neither method is better. The closer the points are to the sides of the triangle, the larger is the statistical difference between methods. Figure 5 shows that the probability that *MixBoost* is better than SWIM for data augmentation is high for all three down-sampled training datasets. For all three plots, the point cloud is concentrated in the bottom right triangle, indicating that *MixBoost* is statistically significantly better than SWIM.

V. Ablation Studies

In this section, we evaluate the specific impact of different components of *MixBoost* and measure the sensitivity of our framework to hyper-parameters.

To balance the consistency of results and ease of analysis and presentation, we restrict our discussion to five datasets, namely Pima Indians, Waveform, PC4, Piechart and Forest Cover. These datasets are selected to ensure maximum *diversity in feature dimensionality* (ranging from 8 features for *Pima Indians* to 54 for *Forest Cover*) and the *extent of class imbalance* (ranging from 1:125 for *Pima Indians* to 1:743 for *Forest Cover*).

For each study (unless otherwise specified), we augment the training dataset with n synthetic hybrid instances, where n is the number of instances in the training dataset. As in main experiments, these synthetic hybrid instances are generated equally over *five MixBoost* iterations ($0.2n$ instances in each iteration). For the scope of this analysis, we use the *EW-Selection* technique in *MixBoost* and sample the interpolation ratio λ from a $Beta(0.5, 0.5)$ distribution.

A. Impact of sampling instances over multiple iterations

In this experiment, we evaluate the importance of generating synthetic hybrid instances over multiple iterations. We define a variant of *MixBoost*, called *MixBoost-1-Iteration* (denoted by *MixBoost-1-Iter*), where all the n synthetic instances are generated in a *single* iteration. This single-step generation is in contrast to *MixBoost* used previously, where the n instances are generated over *five* iterations, $0.2n$ at each iteration.

We augment the training dataset with the generated instances and train the binary classifier on the augmented dataset. Table V shows the results. We observe that the sampling instances over multiple (*five*) iterations in *MixBoost* outperforms the single-step variant on each datasets. This result corroborates the importance of the iterative characteristic ($Boost$ step) of *MixBoost*.

Dataset	***MixBoost-1-Iter***	***MixBoost***
Pima Indians	0.491 ± 0.02	**0.597 ± 0.04**
Waveform	0.843 ± 0.04	**0.844 ± 0.05**
PC4	0.613 ± 0.08	**0.737 ± 0.04**
Piechart	0.721 ± 0.02	**0.741 ± 0.07**
Forest Cover	0.910 ± 0.00	**0.917 ± 0.02**

TABLE V: $g\text{-}mean$ scores for the single step (*MixBoost*-1-Iter) and the proposed *MixBoost*. For all selected datasets, the iterative variant of *MixBoost* outperforms the single step one.

B. Impact of choice of distributions for sampling λ

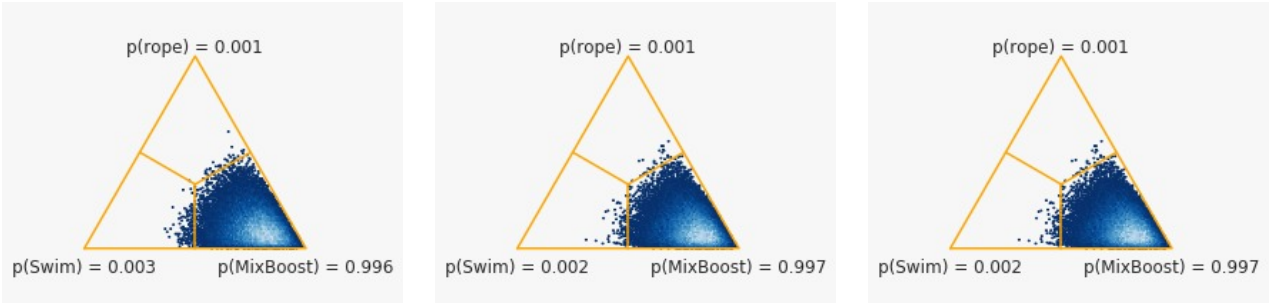
For sampling the interpolation ratio, λ, we experiment with Linear and Non-Linear Probability Density Functions (PDFs). Table VI shows the results. On all datasets, using a non-linear PDF ($Beta(\alpha, \alpha)$ with $\alpha = 0.5$) in the Mix step leads to better classifier performance.

Dataset	**Uniform**	**Beta**
Pima Indians	0.389 ± 0.08	**0.597 ± 0.04**
Waveform	0.661 ± 0.01	**0.844 ± 0.05**
PC4	0.242 ± 0.01	**0.737 ± 0.04**
Piechart	0.312 ± 0.07	**0.741 ± 0.07**
Forest Cover	0.629 ± 0.01	**0.917 ± 0.02**

TABLE VI: $g\text{-}mean$ scores when we sample λ from different distributions. For all datasets, sampling λ from the *Beta(0.5,0.5)* distribution leads to the best performance.

C. Impact of generating hybrid-instances

Existing approaches augment training datasets by over-sampling the minority class instances. *MixBoost* generates synthetic hybrid instances that have elements of both the minority and majority class. Concretely, in the Mix step,



(a) Figure A (b) Figure B (c) Figure C

Fig. 5: Posteriors for SWIM versus *MixBoost* (left and right vertex) on the datasets with 4, 7 and 10 minority class instances ((a), (b), and (c)) for the Bayesian sign-rank test. A higher concentration of points on one of the sides of the triangle shows that a given method has a higher probability of being statistically significantly better [18]. The top vertex *rope* indicates the case where neither method is statistically significantly better than the other.

we generate instances where the target class vector is in-between the majority class vector $[1, 0]$ and the minority class vector $[0, 1]$. Alternatively, in the Mix step, we could have generated synthetic instances where the target class vector corresponded to either the majority or the minority class. To evaluate these alternatives, we create a variant of *MixBoost* in which we assign either the majority class vector (when $\lambda > 0.5$) or the minority class vector (when $\lambda <= 0.5$) to the generated instance. We label this variant *MixBoost-OneHot*. Table VII shows the results comparing *MixBoost* to *MixBoost-OneHot*. We observe that generating hybrid instances with in-between target class labels (from *MixBoost*) outperforms generating instances with either a majority or a minority class label (from *MixBoost-OneHot*). We posit that training with hybrid instances that are not explicitly attributed to either class enables the learning of more robust decision boundaries by helping regulate the confidence of the classifier away from the training distribution.

Dataset	***MixBoost-OneHot***	***MixBoost***
Pima Indians	0.458 $\pm$ 0.03	**0.597 $\pm$ 0.04**
Waveform	0.831 $\pm$ 0.01	**0.844 $\pm$ 0.05**
PC4	0.698 $\pm$ 0.07	**0.737 $\pm$ 0.04**
Piechart	0.566 $\pm$ 0.06	**0.741 $\pm$ 0.07**
Forest Cover	0.898 $\pm$ 0.02	**0.917 $\pm$ 0.02**

TABLE VII: $g\text{-}mean$ scores comparing *MixBoost* to a variant where we generate non-hybrid instances (*MixBoost-OneHot*) in the Mix step. For all considered datasets, generating hybrid instances improves the performance of the binary classifier.

D. Impact of decreasing number of minority class instances

Since *MixBoost* uses an intelligent selection technique for generating synthetic hybrid instances, we want to study whether it outperforms existing approaches with a fewer number of minority class training instances. We generate variations of the training dataset with 2 and 3 minority class instances. We then run *MixBoost* on these variations and train a classifier on the augmented dataset. Table VIII shows the results. *MixBoost* achieves better results than SWIM while using a fewer number of minority class training examples. For instance, for the Waveform and Pima Indians dataset, *MixBoost* outperforms SWIM using half the number of minority class instances (2 vs. 4).

E. Impact of number of generated synthetic hybrid-instances

MixBoost generates synthetic instances iteratively. For each iteration, the classifier is retrained using all instances generated up to this iteration. We expect that the marginal gain of generating instances should decrease as more instances are generated. Figure 6 validates our expectation. We see that the performance of the classifier plateaus as *MixBoost* adds more synthetic instances to the training dataset. This plateau in performance is important for the practical deployment of *MixBoost*.

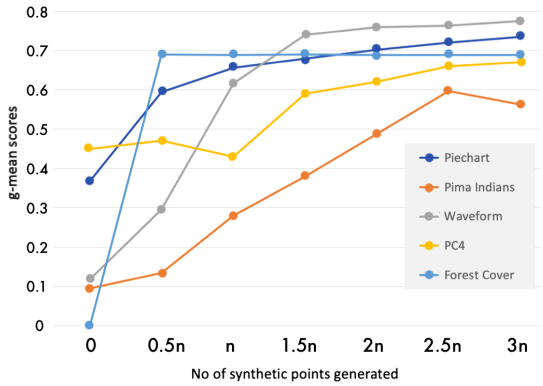


Fig. 6: Variation in $g\text{-}mean$ scores (averaged over 30 runs) for *MixBoost* as we increase the number of generated synthetic hybrid instances. n represents the total number of training instances in the original dataset. The different color lines indicate different datasets. The gain of generating an additional $0.5n$ synthetic instances is initially high and then falls gradually as the number of generated instances increases.

Dataset	SWIM	*MixBoost*		
	$min_{ct} = 4$	$min_{ct} = 2$	$min_{ct} = 3$	$min_{ct} = 4$
Pima Indians	0.499 ± 0.13	0.534 ± 0.08	0.616 ± 0.02	**0.597 ± 0.04**
Waveform	0.652 ± 0.03	0.693 ± 0.06	0.727 ± 0.03	**0.844 ± 0.05**
PC4	0.661 ± 0.06	0.593 ± 0.06	0.688 ± 0.02	**0.737 ± 0.04**
PieChart	0.612 ± 0.05	0.533 ± 0.13	0.579 ± 0.09	**0.741 ± 0.07**
Forest Cover	0.522 ± 0.03	0.452 ± 0.05	0.643 ± 0.03	**0.917 ± 0.02**

TABLE VIII: $g\text{-}mean$ scores for data augmentation using *MixBoost* and SWIM as we reduce the number of minority class instances (min_{ct}) by down-sampling. *MixBoost* achieves best performance on all the datasets. Additionally, *MixBoost* outperforms SWIM on 4/5 datasets when using fewer (3 vs 4) the minority class instances than SWIM.

F. Impact of number of minority class instances

MixBoost mixes an instance of the minority class with an instance of the majority class to create a synthetic hybrid instance. We expect the quality of the generated synthetic instances to improve as we increase the number of minority class instances in the training dataset. The quality of generated instances is measured by classifier performance. Figure 7 shows that the classifier performance improves as we increase the number of minority class instances in the training data.

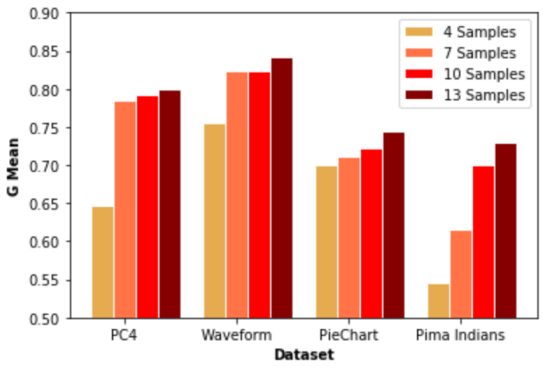


Fig. 7: Variation in $g\text{-}mean$ scores (averaged over 30 runs) for different number of minority class instances using *MixBoost*. For each dataset, the classifier performance increases with an increase in the number of minority class instances.

VI. Related Work

There are broadly two categories of approaches for dealing with classification problems on imbalanced datasets. First, sampling-based approaches and second, cost-based approaches [8], [22]. In this paper, we focus on sampling-based approaches. The most straightforward re-sampling strategies are Random under Sampling (RUS), and Random over Sampling (ROS). They balance the class distribution by either randomly deleting instances of the majority class (RUS) or randomly duplicating instances of the minority class (ROS). However, deleting instances from the training dataset can lead to a loss of information. Further, duplicating instances of the minority class does not add any new information.

SMOTE (Synthetic Minority Oversampling Technique) [5] alleviates these shortcomings. SMOTE balances the class distribution by generating synthetic data instances. SMOTE generates a synthetic instance by interpolating the k-nearest neighbors of a minority class instance in the training data. SMOTE has proved useful in a variety of contexts [5], [6], [12]. However, since SMOTE generates synthetic instances using only minority class instances, the generated instances lie within the convex hull of the minority class distribution. Further, because it does not use majority class instances, SMOTE can increase the overlap between classes. Therefore, if the classes are extremely imbalanced, then SMOTE can degrade classifier performance. Extensions of SMOTE [9] add a post-processing step that tries to remove generated instances that might degrade the performance of the classifier. These methods incorporate the majority class distribution into the post-processing step of the data augmentation process. For instance, Adaptive Synthetic Oversampling (ADASYN) [14], borderline SMOTE [12], and Majority Weighted Minority Oversampling [13] use majority class instances present in the neighborhood of generated instances for post-processing. However, these methods use only the minority class distribution when generating synthetic instances. Therefore, if the number of minority class instances is very small, then the quality of generated instances and subsequently, the performance of the classifier is degraded. [1] create a variant of SMOTE by using the Mahalanobis distance (instead of Euclidean distance) for generating synthetic instances. However, as with SMOTE, they generate instances using only minority class data. SWIM [18] uses information from the majority class to generate synthetic instances and the authors show the technique to outperforms previous methods across several benchmark datasets at extreme levels of class imbalance.

[20] approach the problem of generating synthetic instances from a different direction. They use a Generative Adversarial Network (GAN) trained on minority class data to generate synthetic training instances. However, their method is unable to consistently outperform existing state-of-the-art methods([20], see results). Further, training a GAN for each dataset requires significant compute and hyper-parameter tuning. For these reasons, we have not compared *MixBoost* to their method.

VII. Conclusion

In the present work, we tackle the problem of binary classification on extremely imbalanced datasets. To this end,

we propose *MixBoost*, a technique for synthetic iterative oversampling. *MixBoost* intelligently selects and then combines instances from the majority and minority classes to generate synthetic hybrid instance. We show that *MixBoost* outperforms existing methods on several benchmark datasets. We also evaluate the impact of the different components of *MixBoost* using ablation studies. As directions of future study, we wish to focus on evaluating *MixBoost* for multi-class classification and adapt the idea of iterative sampling and generation through interleaved Mix and $Boost$ steps for regression tasks.

REFERENCES

[1] Lida Abdi and Sattar Hashemi. 2015. To combat multi-class imbalanced problems by means of over-sampling techniques. *IEEE transactions on Knowledge and Data Engineering* 28, 1 (2015), 238–251.

[2] Ali Al-Shahib, Rainer Breitling, and David Gilbert. 2005. Feature selection and the class imbalance problem in predicting protein function from sequence. *Applied Bioinformatics* 4, 3 (2005), 195–203.

[3] Alessio Benavoli, Giorgio Corani, Francesca Mangili, Marco Zaffalon, and Fabrizio Ruggeri. 2014. A Bayesian Wilcoxon signed-rank test based on the Dirichlet process.. In *ICML (JMLR Workshop and Conference Proceedings)*, Vol. 32. JMLR.org, 1026–1034. http://dblp.uni-trier.de/db/conf/icml/icml2014.html#BenavoliCMZR14

[4] Kwabena Ebo Bennin, Jacky Keung, Passakorn Phannachitta, Akito Monden, and Solomon Mensah. 2017. Mahakil: Diversity based oversampling approach to alleviate the class imbalance issue in software defect prediction. *IEEE Transactions on Software Engineering* 44, 6 (2017), 534–550.

[5] Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. 2002. SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research* 16 (2002), 321–357.

[6] Nitesh V Chawla, Aleksandar Lazarevic, Lawrence O Hall, and Kevin W Bowyer. 2003. SMOTEBoost: Improving prediction of the minority class in boosting. In *European conference on principles of data mining and knowledge discovery*. Springer, 107–119.

[7] Dheeru Dua and Casey Graff. 2017. UCI Machine Learning Repository. http://archive.ics.uci.edu/ml

[8] Charles Elkan. 2001. The foundations of cost-sensitive learning. In *International joint conference on artificial intelligence*, Vol. 17. Lawrence Erlbaum Associates Ltd, 973–978.

[9] Alberto Fernández, Salvador Garcia, Francisco Herrera, and Nitesh V Chawla. 2018. SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary. *Journal of artificial intelligence research* 61 (2018), 863–905.

[10] Mikel Galar, Alberto Fernandez, Edurne Barrenechea, Humberto Bustince, and Francisco Herrera. 2011. A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 42, 4 (2011), 463–484.

[11] Salvador García and Francisco Herrera. 2009. Evolutionary undersampling for classification with imbalanced datasets: Proposals and taxonomy. *Evolutionary computation* 17, 3 (2009), 275–306.

[12] Hui Han, Wen-Yuan Wang, and Bing-Huan Mao. 2005a. Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning. In *International conference on intelligent computing*. Springer, 878–887.

[13] H He, Y Bai, E Garcia, and S ADASYN Li. 2008a. Adaptive synthetic sampling approach for imbalanced learning. IEEE International Joint Conference on Neural Networks, 2008 (IEEE World Congress on Computational Intelligence).

[14] Haibo He, Yang Bai, Edwardo A Garcia, and Shutao Li. 2008b. ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*. IEEE, 1322–1328.

[15] Miroslav Kubat, Stan Matwin, et al. 1997a. Addressing the curse of imbalanced training sets: one-sided selection. In *Icml*, Vol. 97. Nashville, USA, 179–186.

[16] S Miri Rostami and M Ahmadzadeh. 2018. Extracting predictor variables to construct breast cancer survivability model with class imbalance problem. *Journal of AI and Data Mining* 6, 2 (2018), 263–276.

[17] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. 2011. Scikit-learn: Machine learning in Python. *Journal of machine learning research* 12, Oct (2011), 2825–2830.

[18] Shiven Sharma, Colin Bellinger, Bartosz Krawczyk, Osmar Zaiane, and Nathalie Japkowicz. 2018. Synthetic oversampling with the majority class: A new perspective on handling extreme imbalance. In *2018 IEEE International Conference on Data Mining (ICDM)*. IEEE, 447–456.

[19] Cecilia Summers and Michael J Dinneen. 2019. Improved mixed-example data augmentation. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 1262–1270.

[20] Fabio Henrique Kiyoiti dos Santos Tanaka and Claus Aranha. 2019. Data Augmentation Using GANs. *arXiv preprint arXiv:1904.09135* (2019).

[21] Yuji Tokozume, Yoshitaka Ushiku, and Tatsuya Harada. 2017. Between-Class Learning for Image Classification. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2017), 5486–5494.

[22] Huihui Wang, Yang Gao, Yinghuan Shi, and Hao Wang. 2016. A fast distributed classification algorithm for large-scale imbalanced data. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*. IEEE, 1251–1256.

[23] Shuo Wang, Leandro L Minku, and Xin Yao. 2014. Resampling-based ensemble methods for online class imbalance learning. *IEEE Transactions on Knowledge and Data Engineering* 27, 5 (2014), 1356–1368.

[24] Hongyi Zhang, Moustapha Cisse, Yann N Dauphin, and David Lopez-Paz. 2017. mixup: Beyond empirical risk minimization. *arXiv preprint arXiv:1710.09412* (2017).

[25] Andrew P. Bradley. 1997. The Use of the Area under the ROC Curve in the Evaluation of Machine Learning Algorithms. *Pattern Recogn.* 30, 7 (July 1997), 1145–1159. https://doi.org/10.1016/S0031-3203(96)00142-2

[26] Shiven Sharma, Anil Somayaji, and Nathalie Japkowicz. 2018b. Learning over subconcepts: Strategies for 1-class classification. *Computational Intelligence* 34, 2 (2018), 440–467.

[27] Diding Suhandy and Meinilwita Yulia. 2017. The use of partial least square regression and spectral data in UV-visible region for quantification of adulteration in Indonesian palm civet coffee. *International journal of food science* 2017 (2017).

[28] Wei Wei, Jinjiu Li, Longbing Cao, Yuming Ou, and Jiahang Chen. 2013. Effective detection of sophisticated online banking fraud on extremely imbalanced data. *World Wide Web* 16, 4 (2013), 449–475.

EasyChair Preprint

№ 4098

Modeling of Wastewater Collection Network Using Arc GIS and SewerGEMS in Kabul, Afghanistan

Ali Reza Noori and S.K. Singh

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

August 27, 2020

Modeling of wastewater collection network using Arc GIS and SewerGEMS in Kabul, Afghanistan

1. Ali Reza Noori, PhD scholar at Delhi Technological University
2. S.K.Singh, Professor and Head Department at Delhi Technological University

Email: a.noori@kpu.edu.af

Abstract

This study was conducted to find out the feasibility of a decentralized wastewater collection system and model it in planned and unplanned areas by using Arc GIS and SewerGEMS software. One of the most critical environmental challenges of the Kabul city is the lack of a sewage collection, treatment, and disposal system. This has caused many health and environmental problems. For this study, the fifth district of Kabul city was selected as a study area and divided into nine study zones. The geometric model of the network was created in Arc GIS using Landuse and landcover, Digital Elevation Model, and Satellite Imagery. Hydraulic analysis and design of the system done by SewerGEMS software and the control of the parameters based on the design results done according to the standard design criteria of sewerage networks. The result of the hydraulic design of the modeled system shows that creating and implementing a decentralized and regional sewer collection network is the best solution for the city. After studying the results of hydraulic analysis and design and after reviewing and controlling the control parameters such as flow velocity, pipe diameter, slopes, pipe profiles, etc., it concluded that the hydraulic model has successfully designed and created. Implementation of such a model is applicable to district scales in the city and looks at the appropriate solution for this issue. It can cover almost all of the planned areas of the city and also most of the unplanned areas.

Keywords: wastewater, sewer networks, Kabul, Arc GIS, SewerGEMS.

1. Introduction

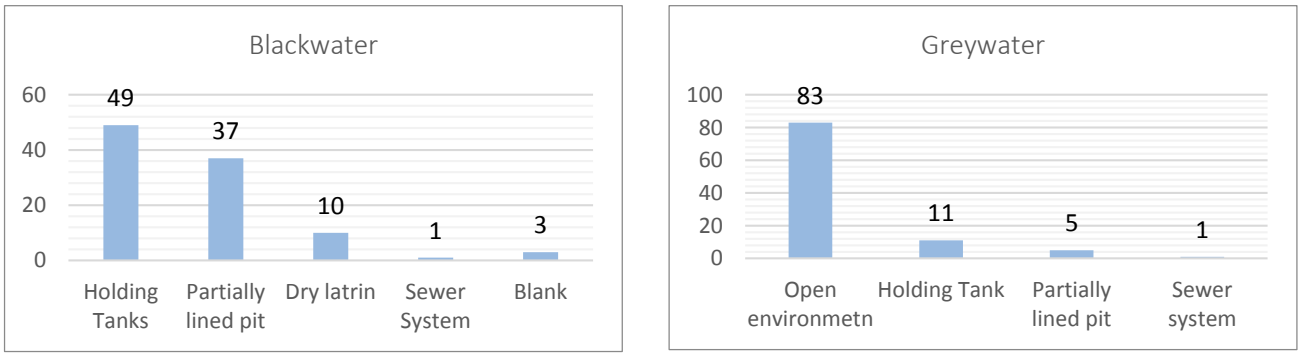
Wastewater collection, in a proper way, as well as its treatment and disposal, is one of the primary hygienic needs of any society. For managers and administrators of urban water and sewage systems, the appropriate solution for different parts of the city is the correct and timely selection of options that must define in terms of physical characteristics, environmental, economic, and social acceptance effects. One of the most critical environmental challenges of the Kabul city is the lack of a sewage collection, treatment, and disposal system. This has caused many health and environmental problems. Currently, Blackwater is conveying sewage tanks, septic tanks, semi-insulated wells, absorption wells, and dry toilets. Sewage from bathrooms, showers, kitchens, etc. (Greywater) is also disposed of using holding tanks, septic tanks, absorption wells, etc., but is mostly discharging into the open environment. The Kabul Municipality has introduced a small scale biological system of the wastewater treatment system and stated that unless severe and immediate action is taken to prevent the contamination of groundwater in Kabul, the groundwater will not be usable for the next few years (shafaqna 2019).

The only valid and available source of drinking water in Kabul is groundwater (Zaryab et al. 2017). Drinking water unfortunately contaminated in Kabul, and most groundwater sources in Kabul unfortunately contaminated with human coliforms (Afzali 2017). Studies show the poor condition of groundwater in Kabul city (Houben et al. 2009; Frahmand 2011; Hayat and Baba 2017; Rhimi and Murakmi 2017; Zaryab et al. 2017; Gesim and Okazaki 2018; Brati et al. 2019; Noori and Nasimi 2019; Zahid et al. 2019). The pollution caused some serious health concerns in the city. Since no official statistics on the disease and its causes are available. According to a report published on the BBC website on 10 Aug of this year(Ali 2017), the leading cause of diseases in Kabul is the use of contaminated water. This news page quotes the doctors of Antani Hospital as saying that the leading cause of epidemics of more than 70 percent of patients who registered in the hospital is drinking contaminated water.

Studies conducted by the Consortium for DEWATS Dissemination Society (2017) in Kabul in districts 2, 3, 4, 5, 10 11, and half of district 17 show that the dispose of black-water and greywater is different. Only one percent of households connected to the sewage collection system; 49 percent of households use holding tanks to dispose of their incoming sewage. 37% use partially lined pit latrines, and 10% use dry toilets. Studies by Paiman and Noori (2019) also show that about 84 percent of households discharge bath and kitchen water directly into the open environment, and 11 percent use the same septic tanks or other different tanks for disposal. Five percent of households use semi-insulated containers. Also, 94% of wells and septic tanks do not have an outlet for overflow flow, but if such a duct is available, it is either connected directly to the open air (37%) or connected to another well (63%) (Figure.1). Conditions for commercial institutions and facilities are slightly different. Most of them (81%) use septic tanks for commodity sewage, 13% use semi-insulated containers, and only 1% use dry toilets. But 1% still use small-scale decentralized treatment plants. Most greywater (59%) is directed directly to the open air. About 26% of these facilities use the same commodity sewage tank or a different tank to transfer this sewage. Five percent of the tanks are semi-insulated, but still, a small number of them have outlets connected to the open space or other wells.

The objective of this study is to find out the feasibility of the wastewater collection system in the Kabul using arc GIS and SewerGEMS software. For this research, a Landuse and Land cover map of the fifth district of Kabul city and Digital Elevation Model (DEM) and satellite imageries used for visualization and geometric modeling. Hydraulic analysis and design of the network done by software tool named Bentley's SewerGEM Version 8i. And parameters control according to the design results done based on the standard design criteria of sewerage networks.

A. Sewage disposal in residential areas



B. Sewage disposal in commercial areas

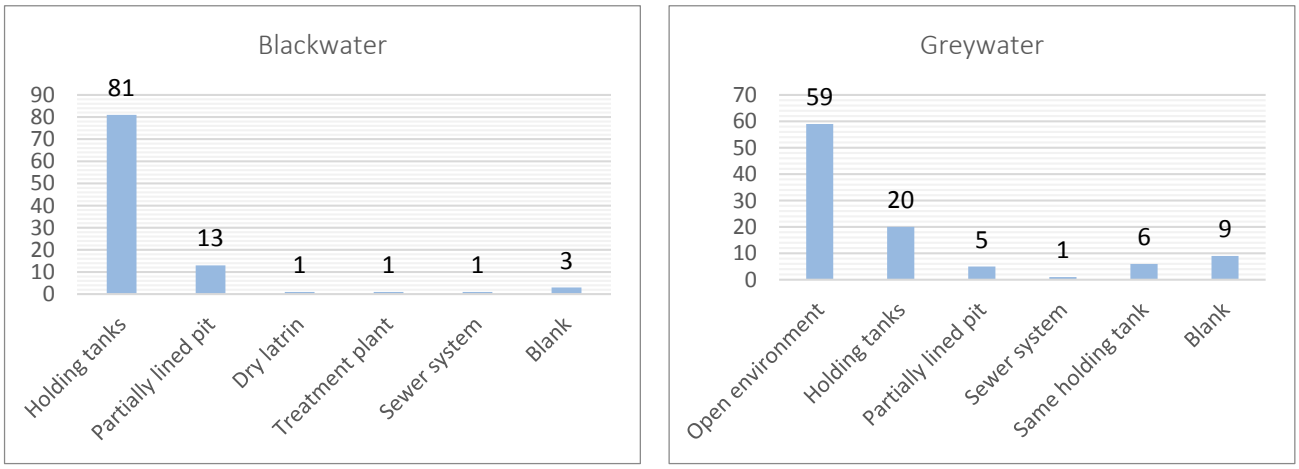


Figure 1. Sewage disposal in Kabul

2. Study area

The process of rapid urbanization is now universal, as more than half of the world's population lives in c ities (Bhave and Rahate 2018). Kabul city is the capital of Afghanistan, located in the central-east of the country. It's population estimated at 4.3 Million (CIA 2020; NSIA 2020). It is the largest national commercial center in the country, and after 2001 a lot of refugees returned to home countries from neighboring countries. Also, a lot of people came to Kabul from other provinces to find a job. The city has 22 municipality districts. District 5 (figure.1) is located in the western part of the town and considered a study area. This district is a populated area of the city, and according to data from Kabul Municipality, it has a population of about 86000. Its total area is about 28 $km^{2,}$ and its perimeter is about 30 km. The reason for selecting this area for the study is that it covers both planned and non-planned area. From the total area of the study area, only about 18% of it has covered by the planned area. The remaining 82% has developed in unplanned mode. Creating models for the unplanned area is time-consuming; therefore, the study area has divided into nine study zones, and the model has developed for only two zones. The total area of this zones area about 9 $km^{2,}$ and one of them covered the planed area, and the other one is an unplanned area (Figure.2). For the selection of the study zones, topographic conditions and future development have considered. If the model developed in the future, it could connect to the main sewer pipes of the existing zones.

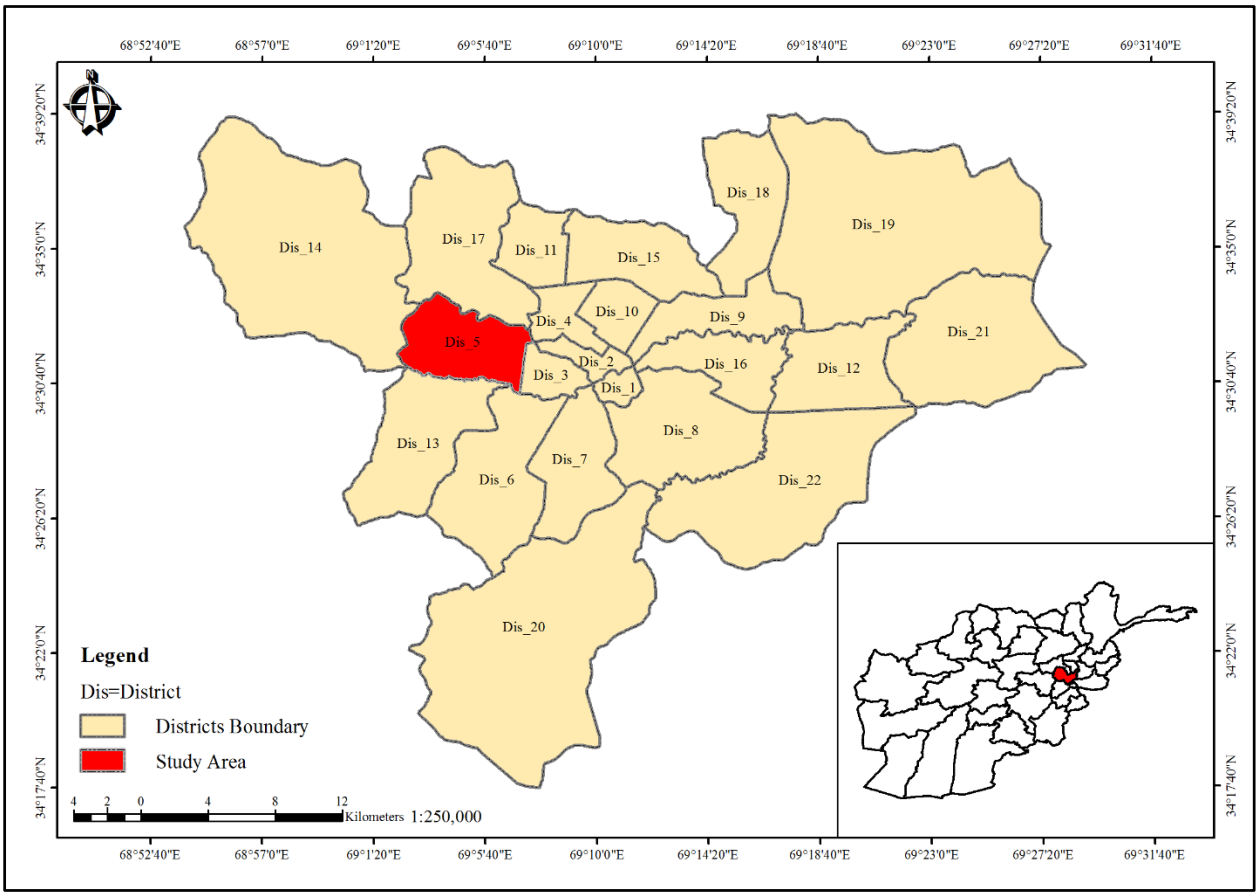


Figure 2. Location of the study area

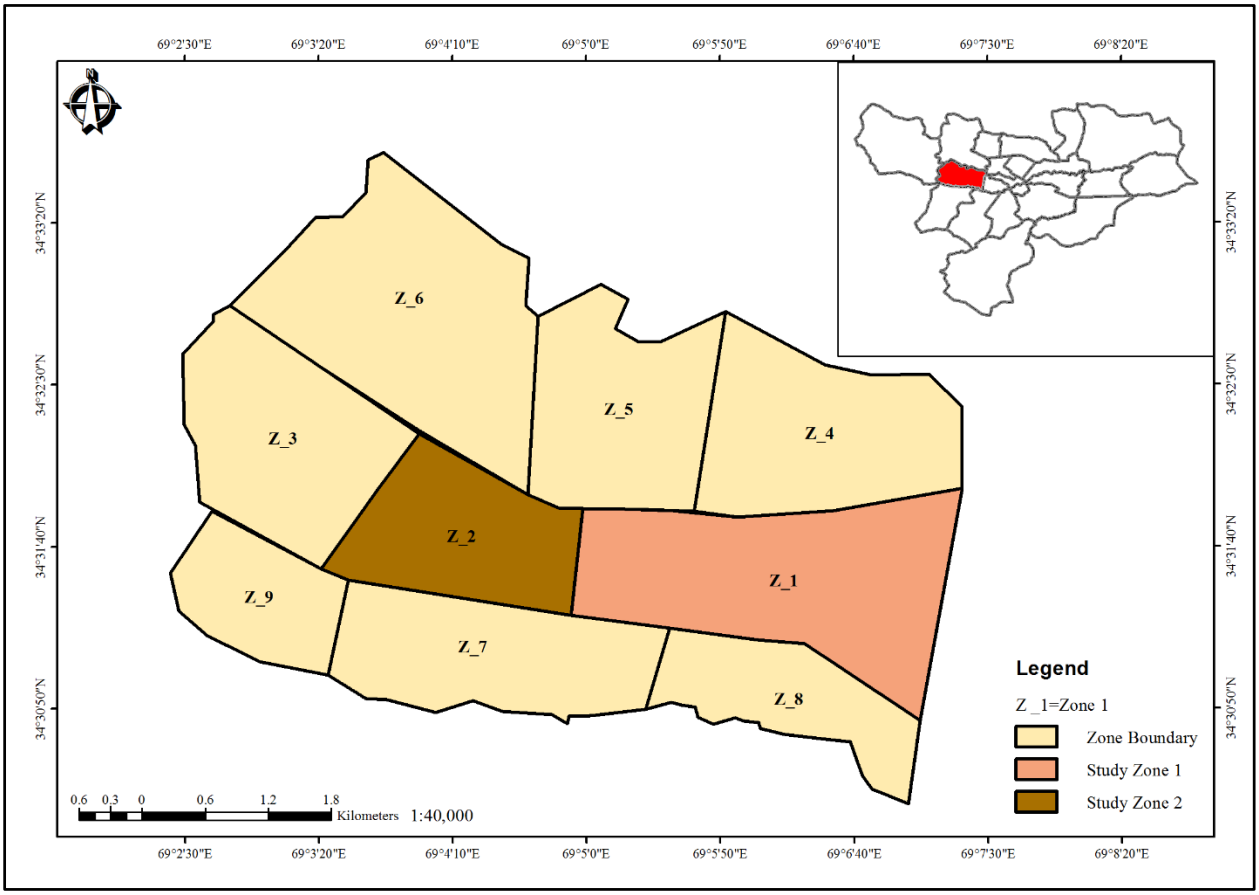


Figure 3. subdivision of the study area.

3. Software application

In this research, the sewage collection network designed in the fifth district of Kabul city. First, the required information for the design, including the length of the sewer pipes, the area of the sewer surfaces and ground elevation at the manholes, has been determined through Arc GIS software. The GIS vastly simplifies the various geospatial scales and can therefore minimize the cost and time of field surveys, which would have a key role to play in environmental conservation (Kumar et al. 2015; Bhave and Rahate 2018). Then, related information like topographic features, statistical information such as population and the per capita water consumption, and wastewater production in the desired geographical area prepared. This information is the basis of the design of the sewage network and developed through reputable sources such as booklets related to geographical knowledge.

To create a geometric model of the system, having land use and landcover is necessary for the study area—Landuse and landcover plan prepared by the ministry of earth and the urban development of Afghanistan. The LULC includes all build-up areas, agricultural, commercial, and institutional areas. The relative location of roads, streets, etc. also has been shown in LULC in the area (figure.3). According to the LULC, aided with satellite images of Kabul city, the number of houses in the area calculated. According to it, the population estimation and flow rate for hydraulic modeling of the system done.

In addition to the LULC aimed proper network routing, there is also a multiplication of contour lines and topographic plans of the area. In practice, it is doing by a geodetic survey and requires a lot of time and budget. However, Digital Elevation Model (DEM) used for research and the creation of basic models. DEM is an open-source data and has different resolutions. In this study, it is used with 30 meters by 30-meter resolutions.

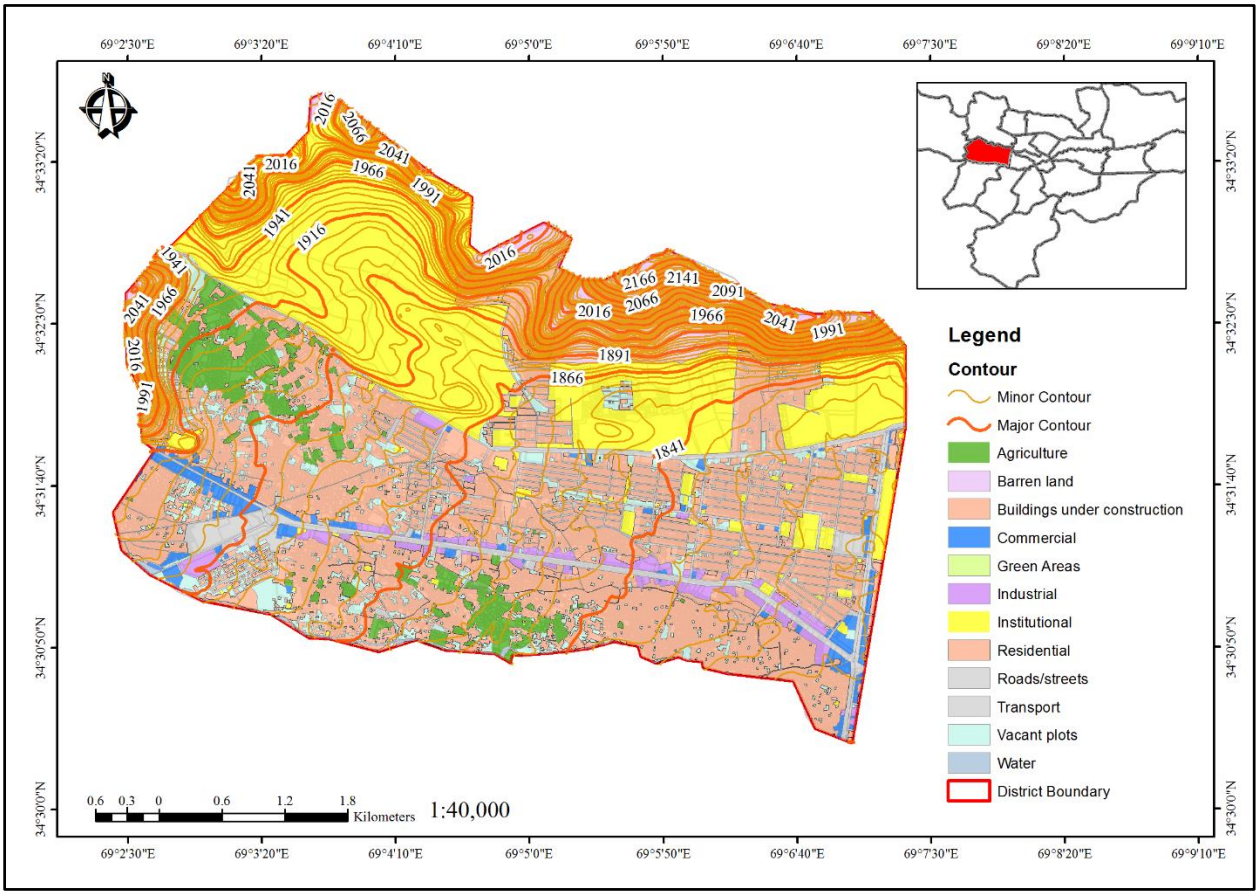


Figure 4. Landuse, Landcover and contour map of the study area

SewerGEMS is one of Bentley's most widely used software and one of the most powerful software for modeling sanitary and hybrid sewage systems. With the help of this software, exact sewage systems can be designed for buildings, houses, and urban context and have many applications in the water and sewage industries. With the help of this software, the municipal sewage system can be designed and optimized. The correct design and construction of the municipal sewage system will significantly help to preserve the environment, the health of humans and other living organisms, and also save water resources. One of the essential features of this software is that it is easy to work with while its powerful features. This software helps water and sewage engineers to be able to design accurate and intelligent systems in a short time and, at the same time, reduce costs. Reduce as much as possible; This software, with its strategic suggestions in various designs, makes the work of engineers very easy and saves their time and energy.

In this research, the essential information and geometric model of the system prepared by Arc GIS 10.6. Then the amount of flow has been received considering the per capita wastewater production for different places, and the required data entered into SewerGEMS for hydraulic design of the network.

4. Geometric modeling

The purpose of creating a geometric model of the system is to hypothetically adjust the routes of the wastewater collection network according to the LULC of the study area and its topography (figure.4). Sewage collection systems must design by the gravitational system so that the sewage flows from the source of production to the treatment unit and is disposed of or reused by gravity. But in many cases, pressure systems can also be used. The geometric model of the system, by default, shows the first figures of the model. In the geometric model, the position of the manholes and primary network routes indicated.

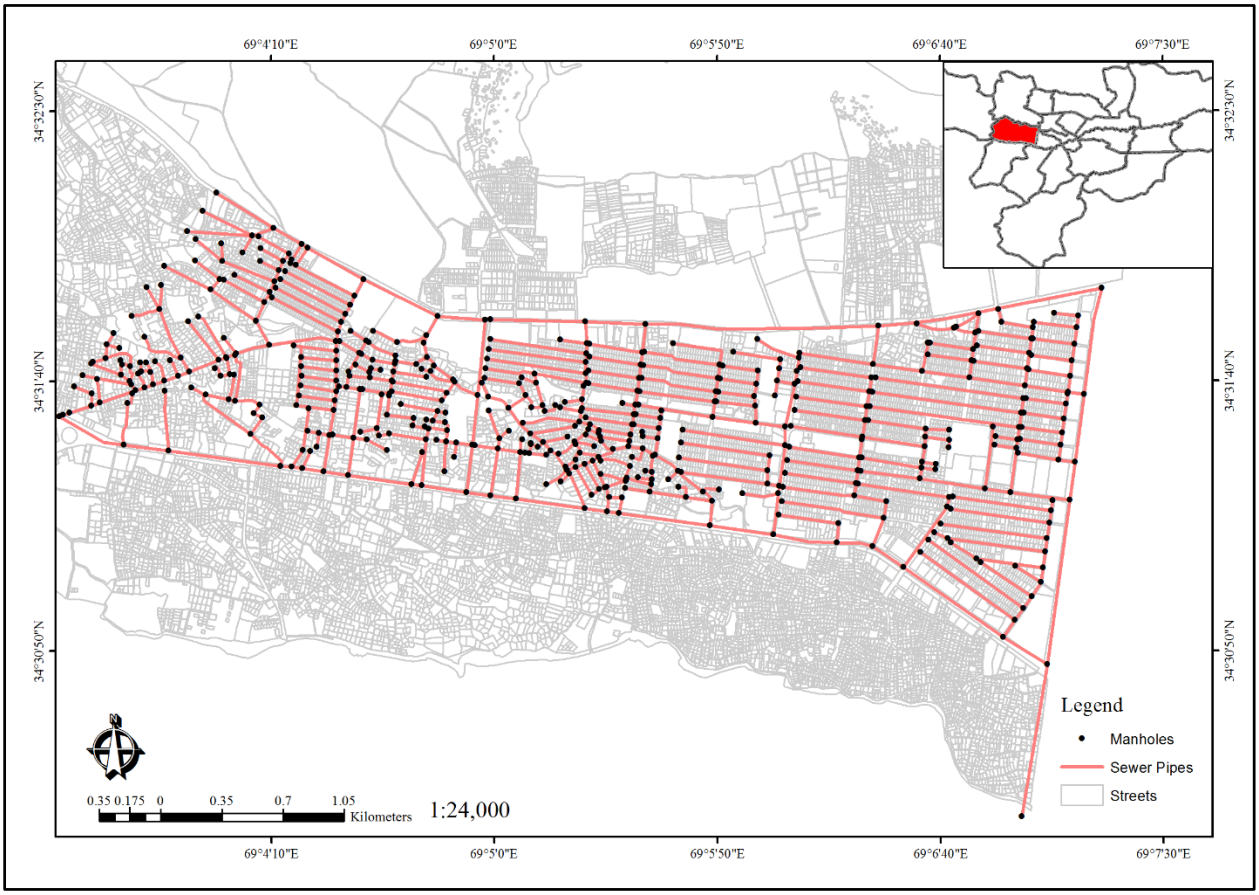


Figure 5. geometric model of the sewer network

5. Result and discussion

5.1. Population density

The total population of Kabul city estimated at 4.3 million and its area is 1023 km^2. The population density is 4500 capita/km^2. The exact area covered (two study zones of the fifth district of Kabul city), which considered in hydraulic design, is 8.834 square kilometers. Therefore, according to the population density, it is possible to find out the population of the covered area, which is about 40,000 people in total. On the other hand, in the area covered, there are mostly medium-sized and high-rise residential blocks, which will have a significant impact on the population density of the city. Therefore, to ensure the presentation of the model and its practical operation, an additional coefficient should be considered for the existing population of the study area. In the next calculations, the reliability coefficient considered being 1.8. Examining the reliability coefficient, the population of the study area is estimated at 72,000 people. Kabul is one of the cities with significant population growth. Therefore, calculating the population increase in hydraulic network design is one of the principles and requirements of accurate and standard design. One of the simplest and most widely used methods of population estimation is the geometric method, in which the following formula used.

$$P_n = P_0(1 + r)^n$$

Where; P_n- Population in the favorable year, P_0 – present population, r – population increase rate, n – years of the design period.

According to the design conditions and considerations, the life of water and sewage projects often considered being about 30 years. The population growth admitted to 2.5 percent per year. Taking into account the above conditions, the population in the next thirty years calculated as follows.

$$P_{30} = 72000\left(1 + \frac{2.5}{100}\right)^{30} = 151025 \text{ capita}$$

5.2. Discharge estimation

The amount of sewage flow considered in the network is the amount of water used for residential purposes. Because this system modeled separately, only the amount of wastewater flow in residential areas without rainwater considered. According to the current situation, the norm of water consumption in the cities of Afghanistan deemed to be 100 liters/capita/day (Seddeqi 2017). On the other hand, the amount of wastewater production generally taken 60-80% of water consumption. But in the current situation in Afghanistan, it is better to consider this norm equal to 100 percent. Because the waste of 20-40% of the water found for irrigation of green areas inside the yard, car washing, etc., this waste is not connecting to the municipal sewage system. But in Afghanistan, because the cost of water is higher and on the other hand, the level of society's economy is lower, and water is not using for car wash purposes inside residential houses as well as in-yard sprinklers, etc. According to per capita demand and a total population of the study area, the total amount of wastewater flow in the covered area can calculate as follows:

$$Q_{avg} = 151025 \times 100 = \frac{15102500\text{lit}}{\text{day}} = 15102\text{m}^3/\text{day}$$

The estimated total flow rate of residential areas in dry climate conditions (Dry Weather flow = DWF) is 15102 cubic meters per day. Based on the design conditions for sewerage networks, additional coefficients should consider ensuring the effective operation of the system. Additional coefficient due to groundwater infiltration, the inflow of rainwater, daily and hourly variation. Taken together, consider all of the above coefficients to be about 2, and the total maximum discharge is calculating as follows:

$$Q_{max} = 15102 \times 2 = 30204\ \text{m}^3/\text{day}$$

5.3. Hydraulic modeling and design

Hydraulic models are the most reliable technique in several types of research on sewer networks (Zhang et al. 2018). The purpose of creating a hydraulic model is to analyze the geometric model, which has already prepared in Arc GIS software. After hydraulic modeling, many of the elements considered necessary in the geometric modeling stage may require to remove from the system. Some of the default paths found in geometric modeling may not meet the hydraulic design requirements of the system. Therefore, hydraulic modeling is count as the central part of the work in the hydraulic design of sewerage networks. To create a hydraulic model, the methodology presented in (Figure.5).

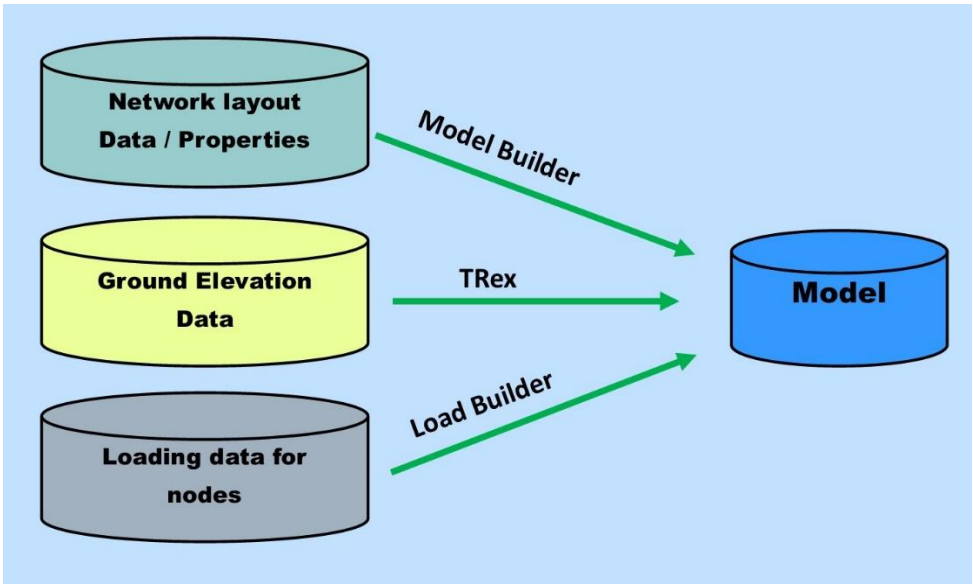


Figure 6. methodology for hydraulic modeling

The purpose of the hydraulic design of the network in the first stage is to find out the required diameter of the sewage pipes, which is determined based on the amount of discharge. Then the calculation of slope and control the velocity of flow in all the sewers. Because velocity is a critical parameter in fluid mechanics and has a specific limit, on the other hand, the sewer pipes are not designing as wholly filled, so the percentage of filling of the pipes should also be determined. In this study, the Area Load Data method used. And the distribution of discharge in the pipes has been done in the form of Equal Flow Distribution. The discharge in all pipes evenly distributed according to the area covered by each pipe. Of course, it should be noted that in fact, such conditions may not appear because some pipes will have more loads, and others will have fewer loads. However, since this model created to study the feasibility of modeling, this method is considered the most appropriate loading method.

The following settings for the hydraulic design of the network applied as a design scenario. The type of pipe material considered for the system is PVC, and its diameter is in the range of 150-600 mm. The maximum and minimum velocity values according to the design criteria, is considered 0.6-3.0 m/sec. The depth of pipes taken with a range of 0.9 – 4.0 m. The slope for the installation of the pipes fixed between 0.005 – 0.1 percent. Flow depth in each pipe considered as 50% full. To evaluate the incoming data and design settings, the created model has validated. After the validation steps, the hydraulic model calculated. The program calculates computational parameters and analyzes analytical parameters.

5.4. Presenting and evaluating the results

The design report provides brief information about the designed model. It includes the necessary details about the designer, such as name, specifications, model title, company, date of modeling, and design. It also provides a summary of the scenario used in model design. It presents the number of elements used in the model, 521 manholes, 521 pipes, and one outfall. Also, it shows the diameter and length of all diameter ranges of the pipes used in the network (Figure7). The total length of the system is 92602.4 meters.

Hydraulic Model Inventory: 5.stsw

Title	مدل سازی شبکه جمع آوری فاضلاب ناحیه پنجم شهر کابل
Engineer	پوهنیار علی رضا نوری
Company	پوهنتون پولی تخنیک کابل
Date	7/26/2019
Notes	

Scenario Summary

ID	1
Label	Base
Notes	
Active Topology	Base Active Topology
User Data Extensions	Base User Data Extensions
Physical	Physical Alternative - 3
Boundary Condition	Base Boundary Condition
Initial Settings	Base Initial Settings
Hydrology	Base Hydrology
Output	Base Output
Infiltration and Inflow	Base Infiltration and Inflow
Rainfall Runoff	Base Rainfall Runoff
Water Quality	Base Water Quality
Sanitary Loading	Base Sanitary Loading
Headloss	Base Headloss
Operational	Base Operational
Design	Base Design
System Flows	Base System Flows
SCADA	Base SCADA
Solver Calculation Options	Base Calculation Options

Network Inventory

Conduits	521	Taps	0
-Circle	521	Transitions	0
-Box	0	Cross Sections	0
-Ellipse	0	Outfalls	1
-Virtual	0	Catchments	0
-Irregular Channel	0	Low Impact Development Controls	0
-Trapezoidal Channel	0	Ponds	0
-Triangular Channel	0	Pond Outlet Structures	0
-Rectangular Channel	0	Headwalls	0
-Pipe-Arch	0	Pumps	0
Laterals	0	Wet Wells	0
Channels	0	Pressure Junctions	0
Gutters	0	SCADA Elements	0
Pressure Pipes	0	Pump Stations	0
Catch Basins	0	Variable Speed Pump Batteries	0
Manholes	521	Air Valves	0

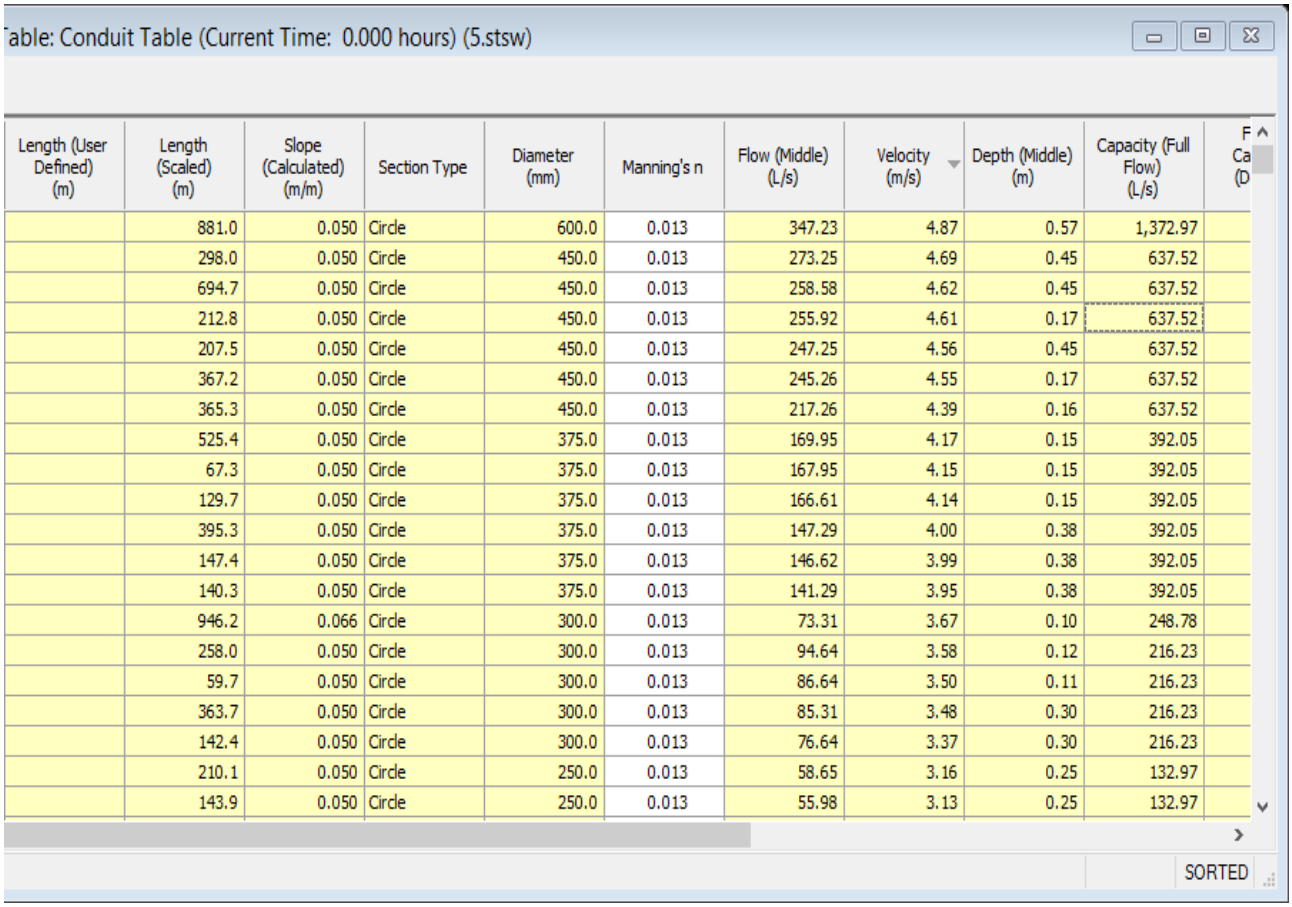
Circle Inventory

Circle - 150.0 mm	80,508.8 m	Circle - 375.0 mm	1,405.3 m
Circle - 200.0 mm	3,732.6 m	Circle - 450.0 mm	2,145.7 m
Circle - 250.0 mm	2,159.0 m	Circle - 600.0 mm	881.0 m
Circle - 300.0 mm	1,770.1 m	Total Length	92,602.4 m

Figure 7. summary of hydraulic design

5.5. Control parameters

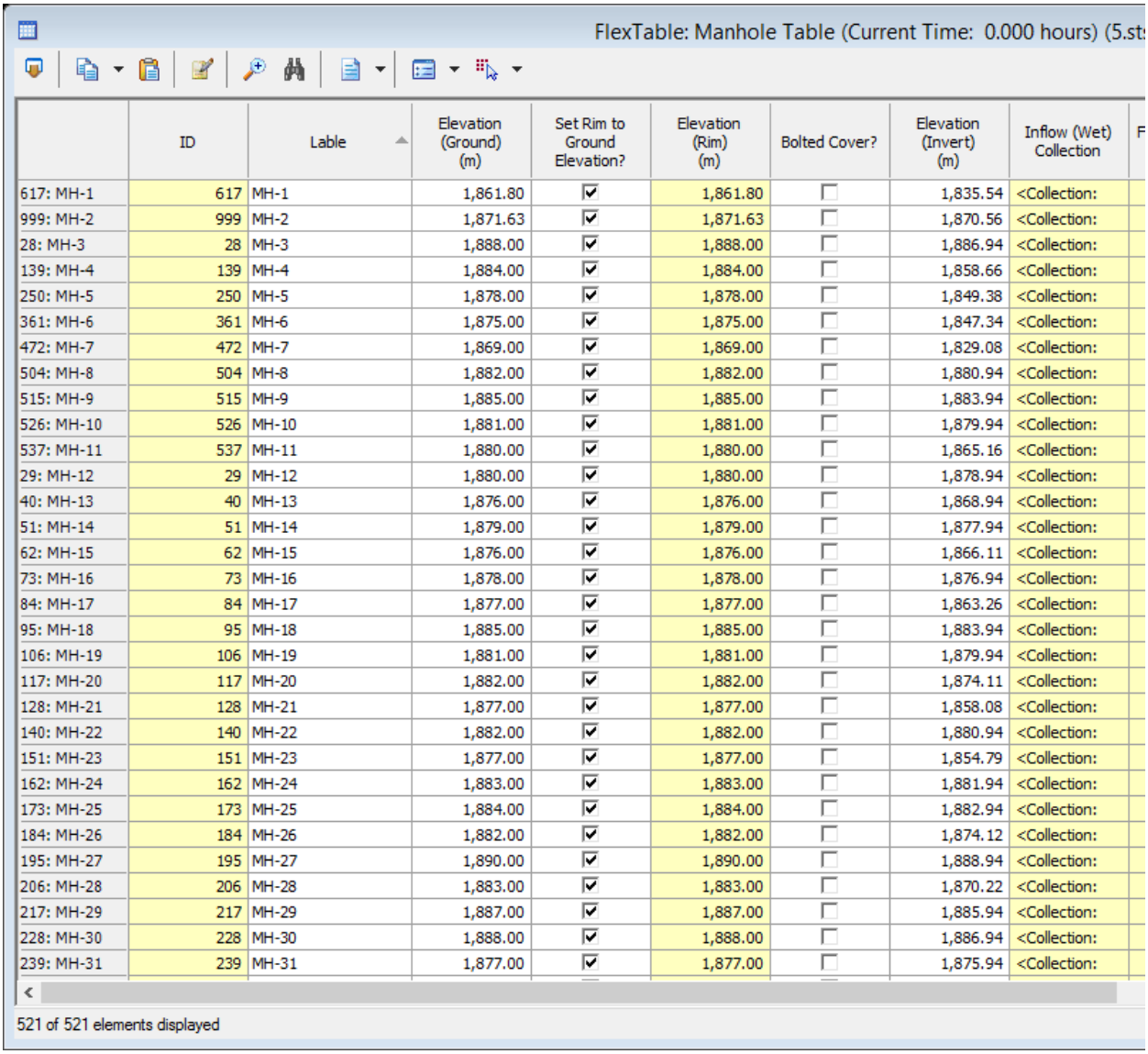
In the design of water and sewage networks, some parameters also identified as control parameters. Because of these computational parameters, the accuracy and correctness of the calculation controlled. For example, velocity is a control parameter that can be calculated and then checked. The velocity limit in sewer networks in the control range is 0.3-5.0 m/sec. The optimum interval is 0.6-1.8 m/sec. Therefore, in modeling and design of sewer networks, the velocity should be within the allowable limit or the desired limit. As illustrated in (Figure.8), the flow velocity is in the range of 0.3-5.0 m/sec. Also, it can control and check other parameters such as pipes diameters that introduced as standard from the program library should be calculated according to the amount of flow. As can be seen in the diameters column, all diameters have been selected based on standard available pipe sizes after calculating the system.



Length (User Defined) (m)	Length (Scaled) (m)	Slope (Calculated) (m/m)	Section Type	Diameter (mm)	Manning's n	Flow (Middle) (L/s)	Velocity (m/s)	Depth (Middle) (m)	Capacity (Full Flow) (L/s)	F Ca (D
	881.0	0.050	Circle	600.0	0.013	347.23	4.87	0.57	1,372.97	
	298.0	0.050	Circle	450.0	0.013	273.25	4.69	0.45	637.52	
	694.7	0.050	Circle	450.0	0.013	258.58	4.62	0.45	637.52	
	212.8	0.050	Circle	450.0	0.013	255.92	4.61	0.17	637.52	
	207.5	0.050	Circle	450.0	0.013	247.25	4.56	0.45	637.52	
	367.2	0.050	Circle	450.0	0.013	245.26	4.55	0.17	637.52	
	365.3	0.050	Circle	450.0	0.013	217.26	4.39	0.16	637.52	
	525.4	0.050	Circle	375.0	0.013	169.95	4.17	0.15	392.05	
	67.3	0.050	Circle	375.0	0.013	167.95	4.15	0.15	392.05	
	129.7	0.050	Circle	375.0	0.013	166.61	4.14	0.15	392.05	
	395.3	0.050	Circle	375.0	0.013	147.29	4.00	0.38	392.05	
	147.4	0.050	Circle	375.0	0.013	146.62	3.99	0.38	392.05	
	140.3	0.050	Circle	375.0	0.013	141.29	3.95	0.38	392.05	
	946.2	0.066	Circle	300.0	0.013	73.31	3.67	0.10	248.78	
	258.0	0.050	Circle	300.0	0.013	94.64	3.58	0.12	216.23	
	59.7	0.050	Circle	300.0	0.013	86.64	3.50	0.11	216.23	
	363.7	0.050	Circle	300.0	0.013	85.31	3.48	0.30	216.23	
	142.4	0.050	Circle	300.0	0.013	76.64	3.37	0.30	216.23	
	210.1	0.050	Circle	250.0	0.013	58.65	3.16	0.25	132.97	
	143.9	0.050	Circle	250.0	0.013	55.98	3.13	0.25	132.97	

Figure 8. Controlling the design parameters

The parameters related to the manholes also can be controlled through SewerGEMS. Among the essential parameters that should consider in manholes are the ground elevation and the depth of the manholes or invert elevation.



	ID	Lable	Elevation (Ground) (m)	Set Rim to Ground Elevation?	Elevation (Rim) (m)	Bolted Cover?	Elevation (Invert) (m)	Inflow (Wet) Collection
617: MH-1	617	MH-1	1,861.80	☑	1,861.80	☐	1,835.54	<Collection:
999: MH-2	999	MH-2	1,871.63	☑	1,871.63	☐	1,870.56	<Collection:
28: MH-3	28	MH-3	1,888.00	☑	1,888.00	☐	1,886.94	<Collection:
139: MH-4	139	MH-4	1,884.00	☑	1,884.00	☐	1,858.66	<Collection:
250: MH-5	250	MH-5	1,878.00	☑	1,878.00	☐	1,849.38	<Collection:
361: MH-6	361	MH-6	1,875.00	☑	1,875.00	☐	1,847.34	<Collection:
472: MH-7	472	MH-7	1,869.00	☑	1,869.00	☐	1,829.08	<Collection:
504: MH-8	504	MH-8	1,882.00	☑	1,882.00	☐	1,880.94	<Collection:
515: MH-9	515	MH-9	1,885.00	☑	1,885.00	☐	1,883.94	<Collection:
526: MH-10	526	MH-10	1,881.00	☑	1,881.00	☐	1,879.94	<Collection:
537: MH-11	537	MH-11	1,880.00	☑	1,880.00	☐	1,865.16	<Collection:
29: MH-12	29	MH-12	1,880.00	☑	1,880.00	☐	1,878.94	<Collection:
40: MH-13	40	MH-13	1,876.00	☑	1,876.00	☐	1,868.94	<Collection:
51: MH-14	51	MH-14	1,879.00	☑	1,879.00	☐	1,877.94	<Collection:
62: MH-15	62	MH-15	1,876.00	☑	1,876.00	☐	1,866.11	<Collection:
73: MH-16	73	MH-16	1,878.00	☑	1,878.00	☐	1,876.94	<Collection:
84: MH-17	84	MH-17	1,877.00	☑	1,877.00	☐	1,863.26	<Collection:
95: MH-18	95	MH-18	1,885.00	☑	1,885.00	☐	1,883.94	<Collection:
106: MH-19	106	MH-19	1,881.00	☑	1,881.00	☐	1,879.94	<Collection:
117: MH-20	117	MH-20	1,882.00	☑	1,882.00	☐	1,874.11	<Collection:
128: MH-21	128	MH-21	1,877.00	☑	1,877.00	☐	1,858.08	<Collection:
140: MH-22	140	MH-22	1,882.00	☑	1,882.00	☐	1,880.94	<Collection:
151: MH-23	151	MH-23	1,877.00	☑	1,877.00	☐	1,854.79	<Collection:
162: MH-24	162	MH-24	1,883.00	☑	1,883.00	☐	1,881.94	<Collection:
173: MH-25	173	MH-25	1,884.00	☑	1,884.00	☐	1,882.94	<Collection:
184: MH-26	184	MH-26	1,882.00	☑	1,882.00	☐	1,874.12	<Collection:
195: MH-27	195	MH-27	1,890.00	☑	1,890.00	☐	1,888.94	<Collection:
206: MH-28	206	MH-28	1,883.00	☑	1,883.00	☐	1,870.22	<Collection:
217: MH-29	217	MH-29	1,887.00	☑	1,887.00	☐	1,885.94	<Collection:
228: MH-30	228	MH-30	1,888.00	☑	1,888.00	☐	1,886.94	<Collection:
239: MH-31	239	MH-31	1,877.00	☑	1,877.00	☐	1,875.94	<Collection:

Figure 9. Ground and Invert Elevation of Manholes

The path profile of the sewer network pipes is another primary parameter that must be controlled. For implementation porpuses of sewage projects, the profile of all pipes required. But here only the farthest path, also called the critical path, is shown (Figure.10& 11). In the profile Ground Elevation, depth of manholes, Invert Elevation, pipe slope, the distance between manholes, and other necessary parameters presented.

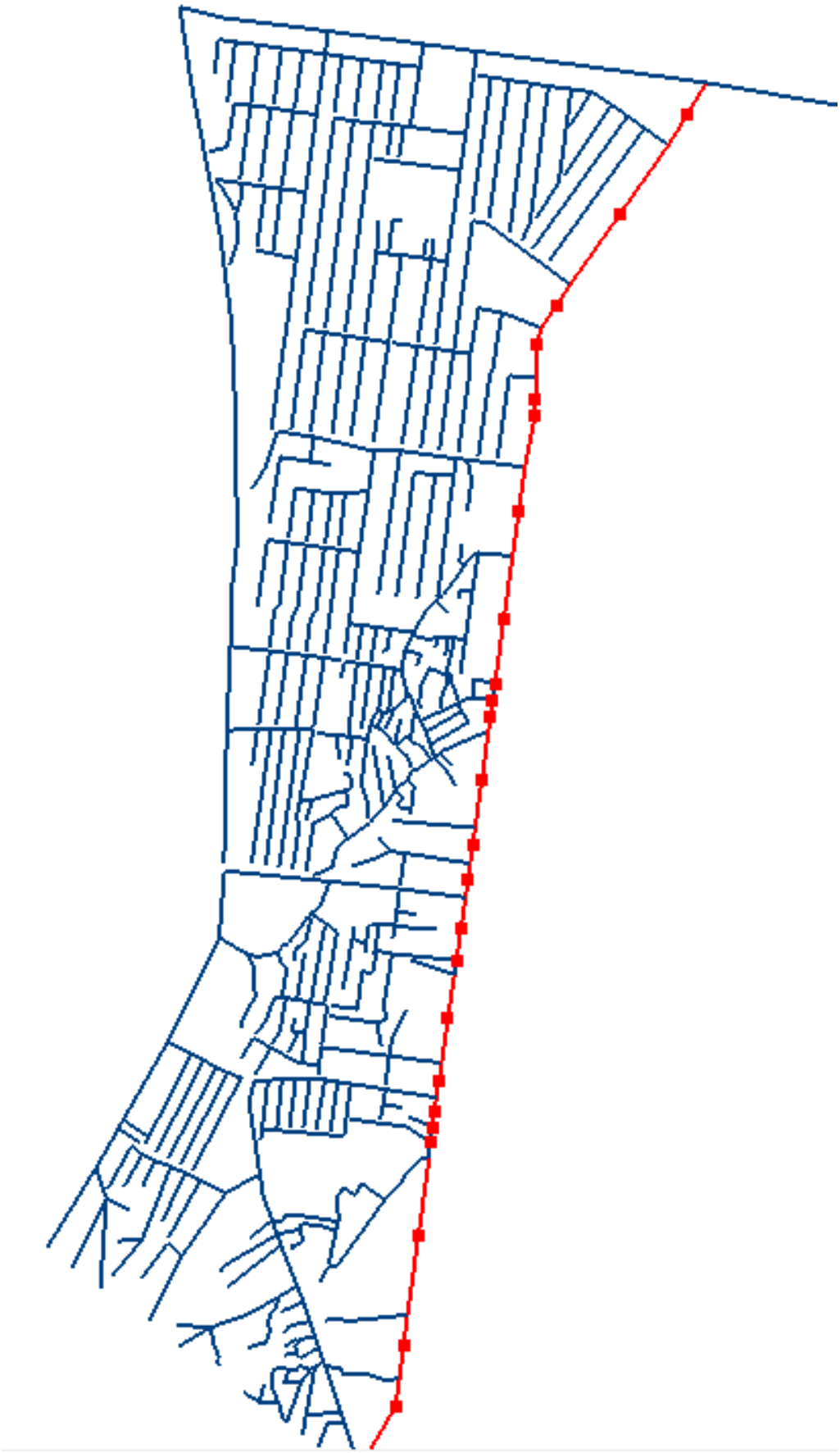


Figure 10. Profile path in plan

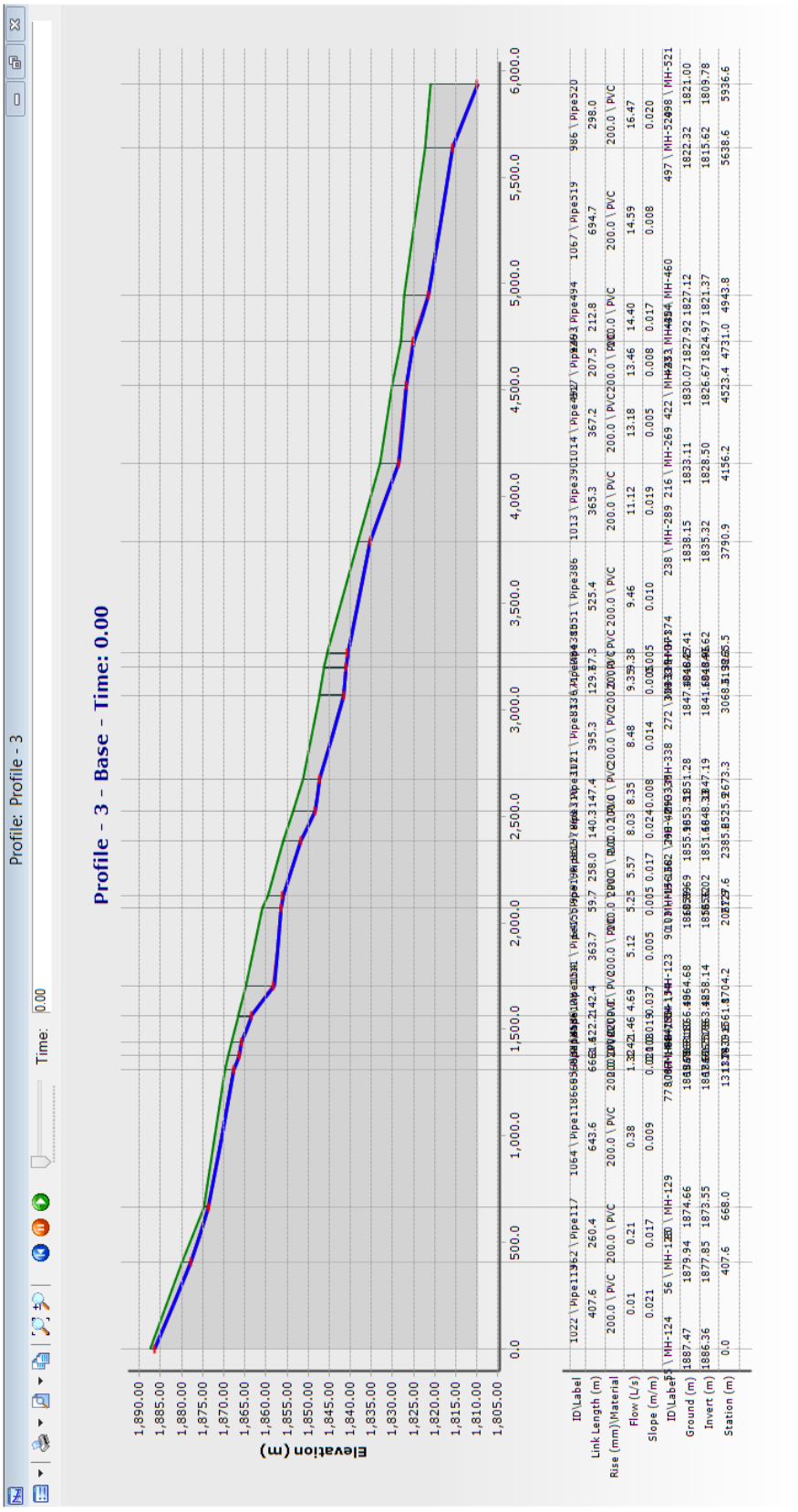


Figure 11. Profile of critical path

5.6. Present research results in the form of graphs and tables

The result of the design can display in different ways. One method already presented in (Figure.7). It is also possible to see manholes, pipes, in detail in the form of separate tables. As well as the pipe material, length, and all required information will show in the program report. The results also can be illustrated in the form of diagrams. By graph presentation, the changes in the parameter in the form of a diagram can be observed in terms

of time in different elements of the model. For example, the variation in the hydraulic gradient in one of the manholes in 24 hours represented (Figure. 12). shows hydraulic gradients in one of the manholes in 24 hours. Of course, this option has a tangible result if the Extended period method used in the design scenario.

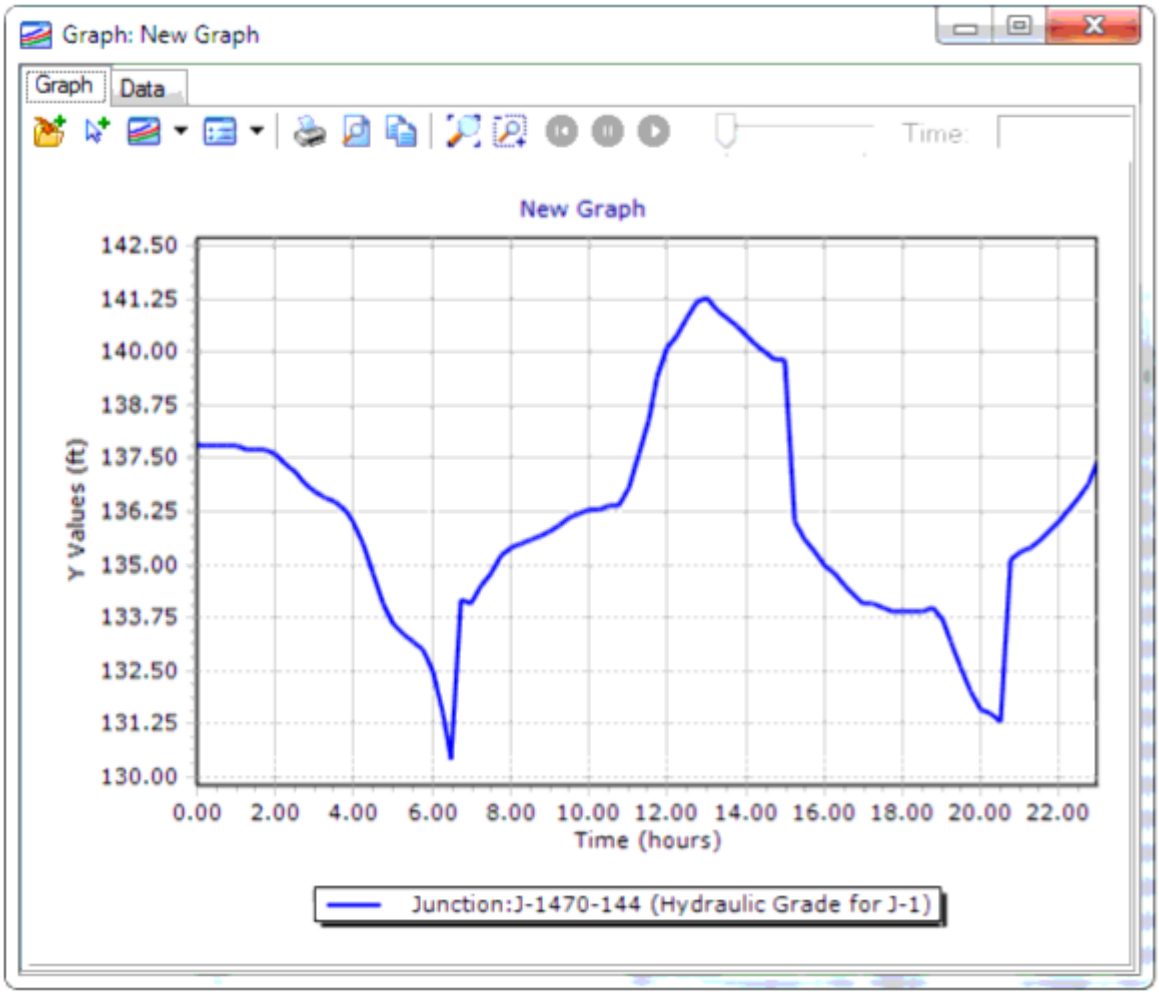


Figure 12. Hydraulic Grade for J-1

6. Conclusion

The main aim of this research was to evaluate the feasibility of decentralized wastewater collection systems or regional systems. Therefore, the fifth region as a sample and two of the nine study zones within it have been studied. LULC and contour map of the study area created using Satellite Image and DEM in the Arc MAP environment. The geometric model also has been created in the Arc MAP environment by creating separate shapefiles for pipes and manholes. The established geometric model is transferred from the Arc GIS environment to the SewerGEMS environment and prepared for hydraulic design.

The hydraulic design of the system was analyzed by SewerGEMS software. As can be seen from the geometric models and hydraulic analysis of the system, it is possible to create a decentralized and regional hydraulic model in Kabul city. Of course, the design of the hydraulic model at the level of the fifth area for two study zones by SewerGEMS software, after creating several models with different scenarios, was finally completed according to the amount of maximum discharge. After interpreting the results of hydraulic analysis and design and after reviewing and controlling the control parameters such as flow velocity, pipe diameter, slopes, pipe profiles, etc., it concluded that the hydraulic model had been successfully designed and created. Such models are applicable to district scales in the city.

As one can see from geometric modeling and satellite images of Kabul city, a very high percentage of Kabul city is unplanned areas. The lack of a regular urban plan causes significant challenges to the creation of such models—this study purpose auxiliary tools in the gravity system for planned and non-planned areas. The use of wet wells and lift stations in the network can mainly solve the system problem.

7. References

Afzali AO (2017) Physicochemical and Bacteriological Quality of groundwater in Kabul basin, Afghanistan. IOSR J Agric Vet Sci 10:58–61. https://doi.org/10.9790/2380-1003025861

Ali H (2017) BBC. In: BBC. https://www.bbc.com/persian/afghanistan-40787693. Accessed 1 Aug 2017

Bhave P, Rahate S (2018) Impact of Redevelopment Projects on Waste Water Infrastructure. J Inst Eng Ser A 99:503–509. https://doi.org/10.1007/s40030-018-0303-x

Brati MQ, Ishihara MI, Higashi O (2019) Groundwater level reduction and pollution in relation to household water management in Kabul, Afghanistan. Sustain Water Resour Manag 5:1315–1325. https://doi.org/10.1007/s40899-019-00312-7

CIA (2020) The World Fact Book. https://www.cia.gov/library/publications/the-world-factbook/geos/af.html. Accessed 30 May 2020

Consortium for DEWATS Dissemination Society (2017) Technical assessment of faecal sludge management in Kabul

Frahmand AS (2011) GIS based Ground Water Quality Analysis in Kabul City Region. Geospatial World Forum 1–13

Gesim NA, Okazaki T (2018) Groundwater Quality Mapping Using Water Quality Index in the Kabul City , Afghanistan. pp 29–34

Hayat E, Baba A (2017) Quality of groundwater resources in Afghanistan. Environ Monit Assess 189:. https://doi.org/10.1007/s10661-017-6032-1

Houben G, Tünnermeier T, Eqrar N, Himmelsbach T (2009) Hydrogeology of the Kabul Basin (Afghanistan), part II: Groundwater geochemistry. Hydrogeol J 17:935–948. https://doi.org/10.1007/s10040-008-0375-1

Kumar A, Mishra RK, Singh SK (2015) GIS Application in Urban Traffic Air Pollution Exposure Study: A Research Review. Suan Sunandha Sci Technol J 2:25–37

Noori KMA, Nasimi MN (2019) Kabul City Groundwater and Need for Artificial Recharge. Proc 4th Int Conf Civil, Struct Transp Eng 1–7. https://doi.org/10.11159/iccste19.215

NSIA (2020) Estimation-Population-Final-Year 2020

Paiman Z, Noori AR (2019) Evaluation of Wastewater Collection and Disposal in Kabul City and Its Environmental Impacts. Mod Environ Sci Eng 5:451–458. https://doi.org/10.15341/mese(2333-2581)/05.05.2019/012

Rhimi O, Murakmi K (2017) Rooftop rainwater harvesting and its efficiency in Kabul new city. 土木学会論文集Ｇ（環境） 73:25–30

Seddeqi MQ (2017) Water Supply Engineering, 2nd edn. Nawesa, Kabul, Afghanistan

shafaqna (2019) shafaqna. In: shafaqna news webpage. https://af.shafaqna.com/FA/341927. Accessed 15 Aug 2020

Zahid S, Asmawi MZ, Abdullah MF (2019) Assessment of the Groundwater Quality and Pollution in Kabul City, Afghanistan. Adv J Tech Vocat Educ 3:20–26. https://doi.org/10.26666/rmp.ajtve.2019.2.4

Zaryab A, Noori AR, Wegerich K, Kløve B (2017) Assessment of water quality and quantity trends in Kabul aquifers with an outline for future drinking water supplies. Cent Asian J Water Res 3:3–11

Zhang D, Martinez N, Lindholm G, Ratnaweera H (2018) Manage Sewer In-Line Storage Control Using Hydraulic Model and Recurrent Neural Network. Water Resour Manag 32:2079–2098. https://doi.org/10.1007/s11269-018-1919-3

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/343657669

Numerical simulations on the aerodynamics of the Ahmed body at different slant angles

Article · August 2020

CITATIONS
0

READS
52

3 authors, including:



Chanfiou Ahmed Mboreha
Nanjing University of Aeronautics & Astronautics
1 PUBLICATION **0** CITATIONS

SEE PROFILE

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

Ventilation strategies in passenger aircraft cabins: An assessment of numerical simulations approaches View project

All content following this page was uploaded by Chanfiou Ahmed Mboreha on 14 August 2020.

The user has requested enhancement of the downloaded file.

Numerical simulations on the aerodynamics of the Ahmed body at different slant angles

Gaurav Kumar[1], Chanfiou Ahmed Mboreha[2] & Yahya Al-fakih[3]

[1]Department of Mechanical Engineering, Delhi Technological University, Delhi, India
Email: gauravdtu@hotmail.com
[2]Department of Aeronautical Engineering, Nanjing University of Aeronautics and Astronautics, Jiangsu, Nanjing, China,
Email: chanfiou2017@nuaa.edu.cn
[3]Department of Mechanical Engineering and Automation, Northeastern University, Shenyang, Liaoning, China, Email: engg.yahya@yahoo.com

Abstract*:* In this research, the aerodynamic behavior of Ahmed's body with slant angles of 0°, 30° and 60° is studied using ANSYS FLUENT CFD code. The flow conditions for all cases were same while the rear slant angle is varied for Ahmed's body. The aerodynamic parameters; lift coefficient and drag coefficient were focused mainly to analyze the flow behavior, it tells the aerodynamics efficiency. The time dependent transient simulation is run to study the behavior of fluid flow. As the Reynold's number lied $5x10^5$in turbulent regime so to accurately capture the turbulent effects, k-epsilon two equation model with Y^+ equal 50 is used. The results of velocity contour, pressure contour, Kinetic Turbulent Energy, velocity vector and velocity profile plots were presented and discussed in detail. The fluid flow behavior illustration of flow separation at rear region of Ahmed's body is captured by defining wake region domain. The results showed that by varying slant angle from $0^{\circ}$to $30^{\circ}$the flow is attached at rear slant region and flow separation is reduced but lift coefficient and drag coefficient values are increased by 4%, and 7% respectively. Whereas increasing the slant angle up to 60°, flow again detached the flow at rear region causing flow separation but there is not much difference in lift and drag coefficients as comparing to 0°slant angle configuration.

***Keywords*:** Ahmed Body, CFD of Ahmed body, numerical solution, separation bubble, C-pillar vortex

Nomenclature:

- φ/θ (degrees) Slant Angle
- S(dimensionless) Strouhal Number
- n(Dimensionless) Frequency of vortex shedding
- D_h(m) Hydraulic diameter of the shape
- V(m/s) Fluid velocity at Re
- Re(Dimensionless) Reynolds number
- P(kg/m^3) Air density
- μ(kg/m.s) Dynamics Viscosity of the air
- L(m) Characteristics linear dimension

I. INTRODUCTION

Fuel consumption is one of most important parameters that are directly related to aerodynamics efficiency of a vehicle [1]. The vehicle efficiency is to be maintained while achieving reduction in fuel

 ©2020, IJA-ERA - All Rights Reserved

consumption by improving external car aerodynamic, CO_2 emissions and noise. Steady increase in power of computers is elevating simulating importance of the vehicle's external aerodynamics and calculating the corresponding characteristic numbers of aerodynamic. A significant part of the growth age of cars is consequently optimization invested. This is achievable particularly and only by development virtually. To create such car industry development processes, a reliable model for simulation validated by fundamental experiments [2].

The experimental studies provide the validated data to make use of bodies of models with characterized geometries by multiple restrictions, such as the back-slant angle. Ahmed body is an example for one such body, a bluff simple shaped body model with basic vehicle aerodynamic properties which is used for multiple investigations reference. The first measurement studied effect of slant angle in range from 0° to 40° at a velocity of flow at 40 m/s ($Re = 5x10^6$) by varying geometry slightly, also called Morel body [3] . The generic Ahmed Body reference model has been chosen as the benchmark for carrying out computations for studying of the aerodynamic parameters. The body was first proposed by Ahmed et al. [4] as can been seen in figure 1, Ahmed et al. conducted measurements under similar boundary conditions, which is called as Ahmed body nowadays. Later, drag forces and surface pressures were determined at full Reynold's number scale by Bayraktar[5] . They used model with geometric size 4.6 times bigger than the Ahmed body original size. Another study on flow behavior of Ahmed body was conducted by Lienhart[6] for measurements on velocity, turbulence, and pressure distribution properties of flow. The objective of that study was to experimentally investigate the Ahmed's body aerodynamic behavior under with different geometrical models and to investigate the CFD analysis using fluent package to compare with experimental information.

An excellent review of different approaches to modeling turbulence for bluff body simulations was given in said research (see references therein) in which it is discussed the use of Reynolds Averaged Navier Stokes (RANS) models, and three more alternatives to be combined with LES simulations: detached eddy simulation (DES), subgrid scale (SGS) models or the more general implicit LES (ILES) method [7].

The body represents the fundamental behavior of aerodynamics of a vehicle, rear part especially. The slant angle ϕ has a high impact on the aerodynamic lift and drag at the rear part. Both coefficients of lift and drag change slightly at the angle ϕ=30° when moved from ϕ=0. Thus, it is named as critical slant angle

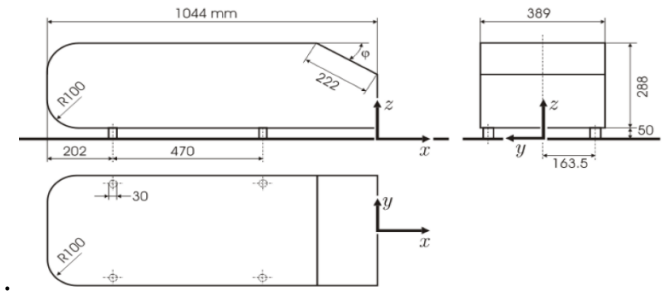


Figure 1: Ahmed's body dimensions sample [8]

In the automotive industry the simulation of 3D CFD flows through the vehicle is an essential tool for understanding the flow phenomenon for aerodynamics design. The flow behavior to the drag force of the body is complex phenomenon to be captured by turbulent models. Some important features of the

 ©2020, IJA-ERA - All Rights Reserved

flow contributing in the fluid flow is ground flow of the Ahmed body and the separation of the flow from the vehicle, which is essential to be captured for the accuracy of the drag predictions [9].

II. METHODOLOGY

A. CFD Flow chart

The CFD process flowchart for the analysis of Ahmed's body is shown below:

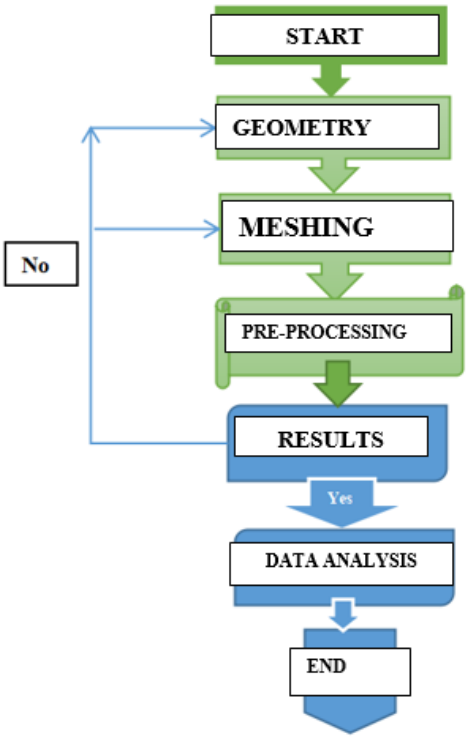


Figure 2: Flow-chart

B. Mathematical Formulations:

The mathematical equations that has been used in this report are listed below

i. Pythagorean Theorem:

The Pythagorean Theorem was applied to find sides length of Ahmed body using the available angles and vice versa as given in equation (1). This theorem is applicable to find length of one side by using the other two lengths of the right-angle triangle as given in equation (2) and illustrated in figure (3):

Equation 1: Pythagorean Theorem Formula for angles.

$$\theta = \sin^{-1}\left(\frac{\text{opposite}}{\text{Hypotenuse}}\right)$$

©2020, IJA-ERA - All Rights Reserved

Equation 2: Pythagorean Theorem Formula for side length.

$$(Hypotenuse)^2 = (opposite)^2 + (Adjacent)^2$$

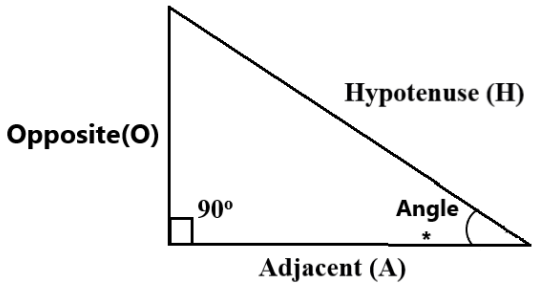


Figure 3: Illustration of Pythagorean Theorem on a triangle.

Area of Rectangle:

The formula to obtain area of rectangle by multiplication of length to width as given in equation (3)

A=LW; A=Area, L=Length, W=width ……(3)

ii. Frequency of shedding:

The frequency of shedding formula as given in equation (4).

$$S = \frac{n\,D}{V} \qquad \text{…..(4)}$$

iii. Vortex shedding period:

The formula to find the vortex shedding period as given in equation (5)

$$T = \frac{1}{n} \qquad \text{…(5)}$$

iv. Reynolds number:

The formulation to obtain the Reynolds number is given in equation (6) as it signifies the flow field behavior whether it is laminar or turbulent.

$$Re = \frac{\rho\, V\, L}{\mu} \qquad \text{…..(6)}$$

v. 3D Hydraulic diameter for Rectangle:

The mathematical formulation to find hydraulic diameter for rectangle is given in equation (7)

$$D_h = \frac{2\ x\ length\ x\ height}{Length + height} \qquad \text{….(7)}$$

©2020, IJA-ERA - All Rights Reserved

C. Geometry of Ahmed's body:

For defining slant angles on Ahmed's body to create 30° and 60° slant angles at back surface, Chamfer button was selected to create this sharp edge. As those two cases are different in their angles, this step had different right and left length sides values. For determining those values for each of slant angle cases, two mathematical equations are used: equation (1) and (2). Ahmed's body shapes with three different angles (0°, 30°, 60°) are presented in figure 4

For 30° angle:

$$30 = \sin^{-1}\left(\frac{\text{Right length}}{222}\right) \rightarrow \text{Right length} = 111\ mm$$

$$(222)^2 = (111)^2 + (Left\ length)^2 \rightarrow Left\ length = 192.26\ mm$$

For 60° angle:

$$60 = \sin^{-1}\left(\frac{\text{Right length}}{222}\right) \rightarrow \text{Right length} = 192.26\ mm$$

$$(222)^2 = (192.26)^2 + (Left\ length)^2 \rightarrow Left\ length = 111\ mm$$

a. Angle=0° b. Angle=30° c. Angle=60°



Figure 4: Ahmed's body shapes with three different angles (0°, 30°, 60°)

D. Meshing generation and Boundary conditions in the three cases

i. Meshing generation

The most important part of CFD simulation is the mesh generation, as the results of CFD parameters: specific pressure and stress are calculated on these mesh cells defined by meshing. The division of the domain into small elements is done to simulate CFD solution with good convergence. The following steps were adopted for mesh generation, firstly the Reynolds number ($5x10^5$) is utilized to measure the flux velocity that is 25.363m/s. For defining inflation layer thickness layers on Ahmed's body, Y^+ method is used to define inflation layer thickness values for capturing wall effects near Ahmed's body surfaces. By knowing that the value of Y^+ close to 30 is more desirable for wall functions so y+ has been taken to be 50 for getting more accurate results with all the air properties.

There are three refinement regions in Ahmed's body cases as defined below and as shown in figure 5:

First region – Transition from far-field to near field. Second region – Underbody of the body. Third region – Wake region

 ©2020, IJA-ERA - All Rights Reserved

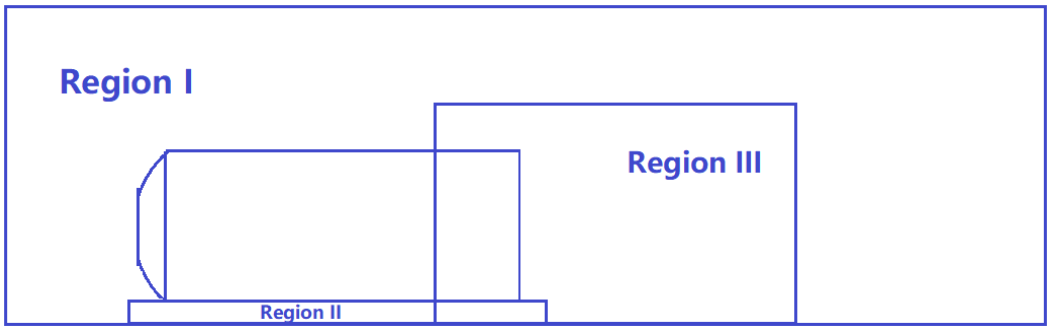


Figure 5: Regions mesh refinements.

ii. Results of the meshing parameters and mesh distribution:

The mesh generated on whole body is visualized here, the meshing parameters and distribution are quite refined and are reliable as compared with external sources. The mesh results are shown in figure .

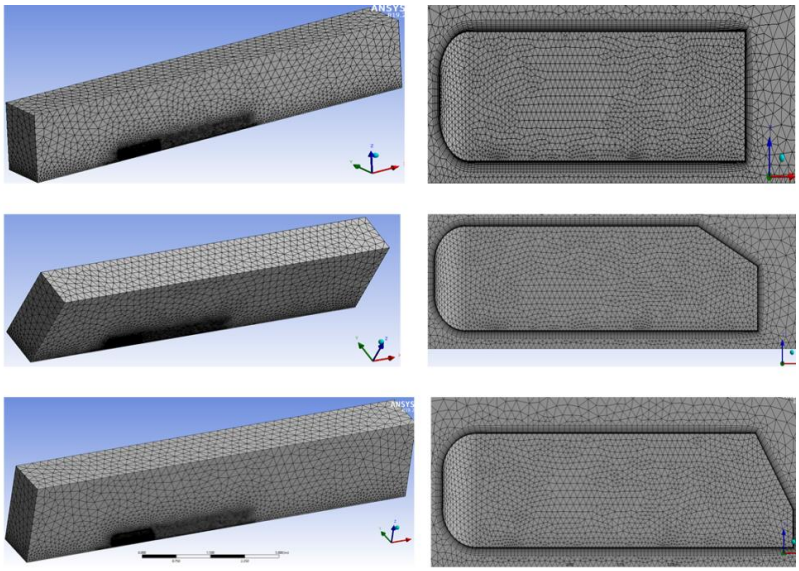
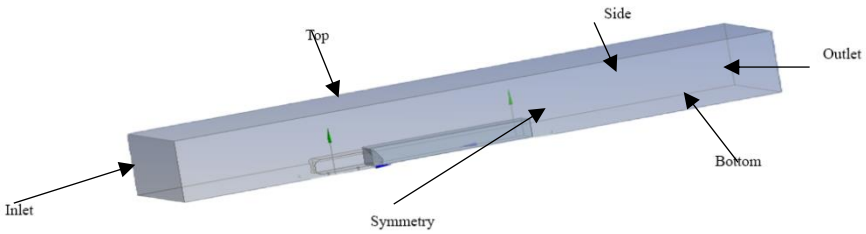


Figure 6: End of meshing of domain and Ahmed's body at 0°, 30° and 60°

iii. Boundary conditions in the three cases

All boundary conditions in the three cases are illustrated in the figure 7.



 ©2020, IJA-ERA - All Rights Reserved

Figure 7: Boundary conditions in the three cases of Ahmed's body.

Table 1: Summary of the Mesh Parameters and mesh distribution

Parameter	Ahmed's body at 0° slant angle	Ahmed's body at 30° slant angle	Ahmed's body at 60° slant angle
Number of elements	555029	547085	553100
Number of nodes	156289	145666	155102
Element size	0.2 m	0.2 m	0.2 m
First layer height	0.00064 m	0.00064 m	0.00064 m
Maximum layers	12	11	11

E. Numerical Method:

The mesh has been transferred to FLUENT setup for defining boundary conditions and pre-processing for solving the three cases of Ahmed's body. The transient flow simulation is conducted for all three cases of Ahmed's body CFD analysis using pressure-based solver and the effects of gravity are taken in consideration. In the derivation of the k – epsilon model, the assumption is that the flow is fully turbulent, and the effects of molecular viscosity are negligible. The standard k –epsilon model is therefore valid only for fully turbulent flow[10]. The pressure-based solver along with SIMPLE algorithm is adopted under solution schemes. For the turbulent kinetic energy specific dissipation rate, momentum and pressure equations, 2nd order upwind discretization is selected. The gradient calculation from least square cell based is chosen. The solution solver setup with solver schemes are tabulated below for all the slant angles configurations of Ahmed's body. Here, we used reference area value as 0.11205 m^2 and velocity of air is 25.363 m/s with Pressure based solver is used because Mach number is low hence we can assume flow as incompressible and as due to capture flow at time intervals transient case is preferred.

Calculations that have been done in table 4:

Physical properties of Air that has been used are shown in table 5, these assumptions were considered for CFD analysis:

1. The surface of the Ahmed body was modeled using walls with no-slip.
2. The condition of the flow did not involve the temperature or the energy equation.
3. Incompressible flow.

Table 2: Air properties used in simulation.

Parameters	Values
Density	1.225 Kg/m3
Viscosity	1.7894x10-5 kg/m s
Ratio of specific heat	1.4
Temperature	288.16 K

At Re=$5x\ 10^5$:

 ©2020, IJA-ERA - All Rights Reserved

By known that the linear dimension = 0.288 m, the Velocity can be calculated by using the equation below:

$$Re = \frac{\rho\, V\, L}{\mu} \rightarrow 5\text{x } 10^5 = \frac{1.225\ x\, V\, x\, 0.288}{1.7894\text{x}10 - 5} \rightarrow V = 25.35998 \frac{m}{s}$$

The D_h for Ahmed's body inlet can be determined from the equation 7:

$$D_h = \frac{2\ x\ length\ x\ height}{Length + height} = \frac{2\ x\ 0.389\ x\ 0.288}{0.389 + 0.288} = 0.331 \text{ m}$$

1. From S vs. Re chart:

$S \approx 0.2$ At $Re = 5x\ 10^5$

2. Frequency of shedding:

By knowing that the velocity is 25.363 m/s at $Re = 5x\ 10^5$ and$D_h = 0.331\ m$, Frequency of vortex shedding can be calculated using equation 4:

$$S = \frac{n\ D}{V} \rightarrow n = \frac{25.363 \text{ x } 0.2}{0.331}$$

$n = 15.325$

3. Vortex shedding period:

By knowing the value of Frequency of vortex shedding, period can be calculated using equation 5:

$$T = \frac{1}{n} = \frac{1}{15.32507} = 0.0653 \text{ s}$$

4. Time step size:

By break the period into 50 pieces

$$\Delta T = \frac{\text{Vortex shedding period}}{50} = \frac{0.0653}{50} \approx 0.001305$$

5. Number of time steps:

$$N = \frac{Total\ time\ for\ 20\ period}{\text{Time step size}} = \frac{0.0653\ \text{x } 20}{0.001306} \approx 1000$$

6. Area of the Ahmed's body inlet

$$Area = Length\ x\ Width = 0.389\ x\ 0.288 = 0.112032\ m^2$$

 ©2020, IJA-ERA - All Rights Reserved

III. SIMULATION OUTCOMES

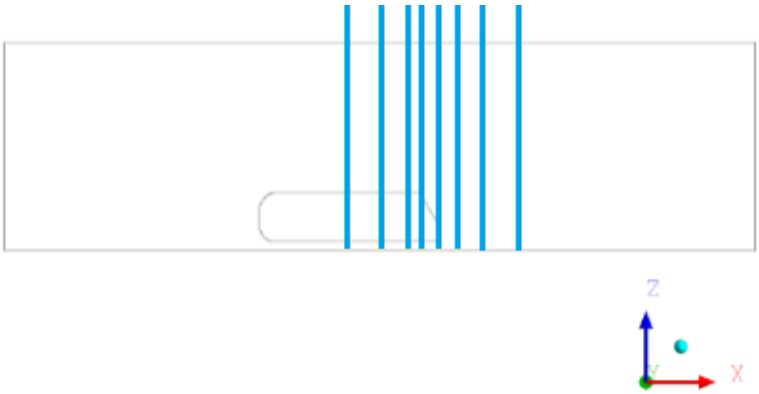


Figure 8: Illustrating the location of the lines on Ahmed's body for creating velocity profile

As the CFD simulation of Ahmed's body is completed at defined inputs and convergence criteria so the results are obtained using post processing through CFD Post. The velocity contours, velocity vector, pressure contours and turbulence kinetic energy contours are obtained as results to analyze the performance of Ahmed's body. At the end (Insert → Location → Line) was selected to create the velocity profile charts for each of the slant angle Ahmed's body case. The following steps in table (6) has been followed for all the slant angles cases in order to capture the flow over at the rear side of the Ahmed's body and monitor lines to be formed in two positions as illustrated in figure (8).

Table 3: Inserting lines to create velocity profile for all slant angles.

Line 1			
Point 1	-0.544	0	1.5
Point 2	-0.544	0	0
Line 2			
Point 1	-0.344	0	1.5
Point 2	-0.344	0	0
Line 3			
Point 1	-0.194	0	1.5
Point 2	-0.194	0	0
Line 4			
Point 1	-0.094	0	1.5
Point 2	-0.094	0	0
Line 5			
Point 1	0	0	1.5
Point 2	0	0	0
Line 6			
Point 1	0.106	0	1.5
Point 2	0.106	0	0
Line 7			
Point 1	0.256	0	1.5
Point 2	0.256	0	0

©2020, IJA-ERA - All Rights Reserved

Line 8			
Point 1	0.456	0	1.5
Point 2	0.456	0	0

A. Lift coefficient and Drag coefficient profiles for the three cases of Ahmed's body.

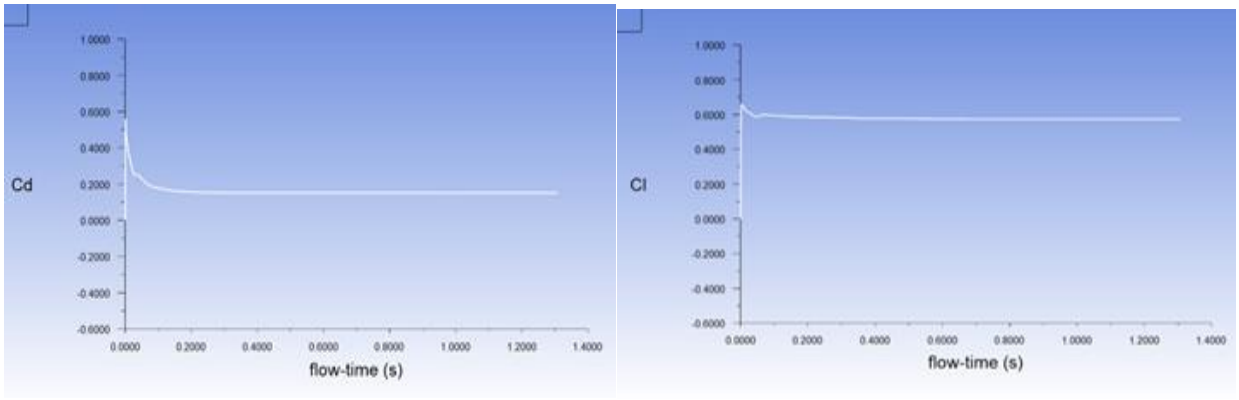


Figure 9: Lift coefficient (Right) and Drag coefficient (left) for 0°slant angle

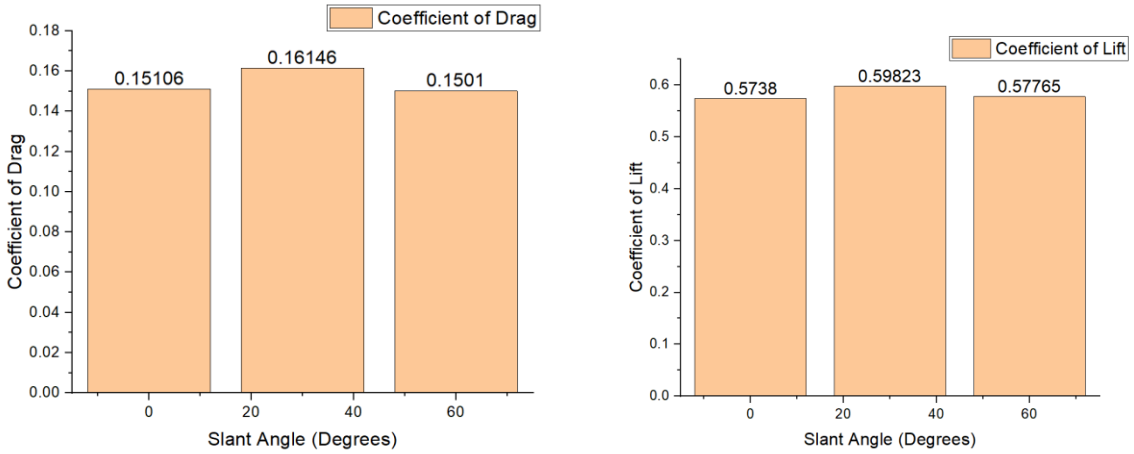


Figure 10: Coefficient of Drag vs Slant Angle & Coefficient of Lift vs slant angle

i. Lift coefficient and Drag coefficient for **0**°*slant angle:*

Lift coefficient and Drag coefficient plots illustrated that the drag and the lift values converged to 0.15016 for Cd and 0.57380 for Cl for Ahmed's body with 0° slant angle as shown in figure 9. The lift coefficient and drag coefficient profiles are showing straight line while converging to final iteration shows that values of lift coefficient and drag coefficient are quite accurate for defined flow condition.

i. Lift coefficient and Drag coefficient plot for **30**°*slant angle:*

Lift coefficient and Drag coefficient plots illustrated that the drag and the lift values converged to 0.16146 for Cd and 0.59823 for Cl for Ahmed's body with 30° slant angle as can be seen in figure 11. The lift coefficient and drag coefficient plots are showing straight line while converging to final

 ©2020, IJA-ERA - All Rights Reserved

iteration shows that values of lift coefficient and drag coefficient are quite accurate for defined flow conditions.

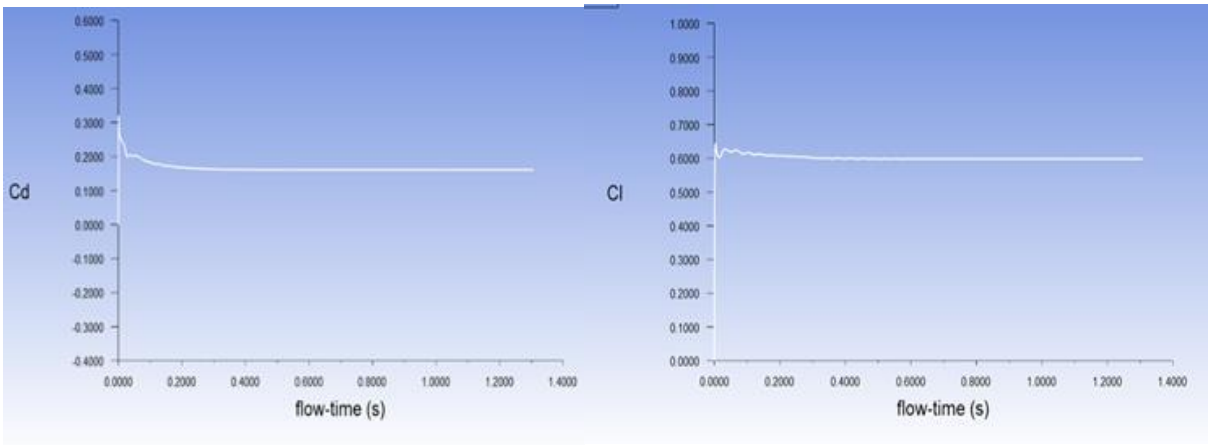


Figure 11: Lift coefficient (Right) and Drag coefficient (left) for 30°slant angle

ii. Lift coefficient and Drag coefficient for **60**°*slant angle:*

Lift coefficient and Drag coefficient plots illustrated that the drag and the lift values converged to 0.15010 for Cd and 0.57763 for Cl for Ahmed's body with 6° slant angle as shown in figure 12. The lift coefficient and drag coefficient plots are showing straight line while converging to final iteration shows that values of lift coefficient and drag coefficient are quite accurate for defined flow conditions.

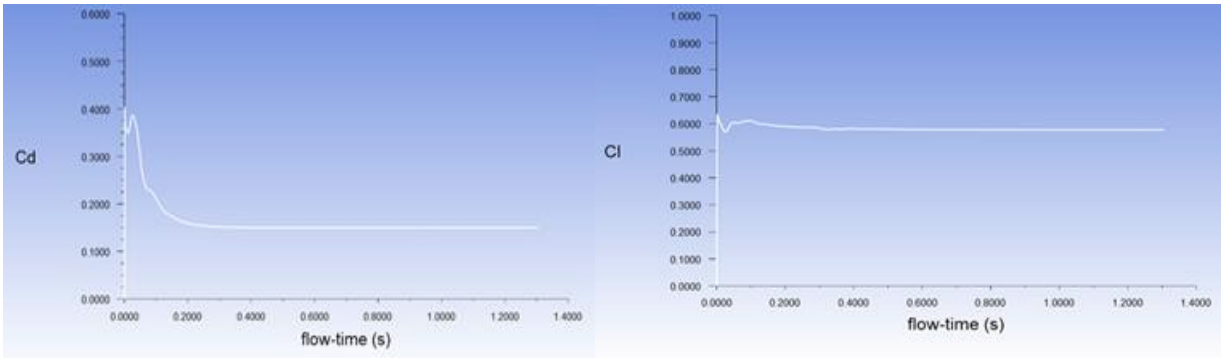


Figure 12: Lift coefficient (Right) and Drag coefficient (left) for 60°slant angle

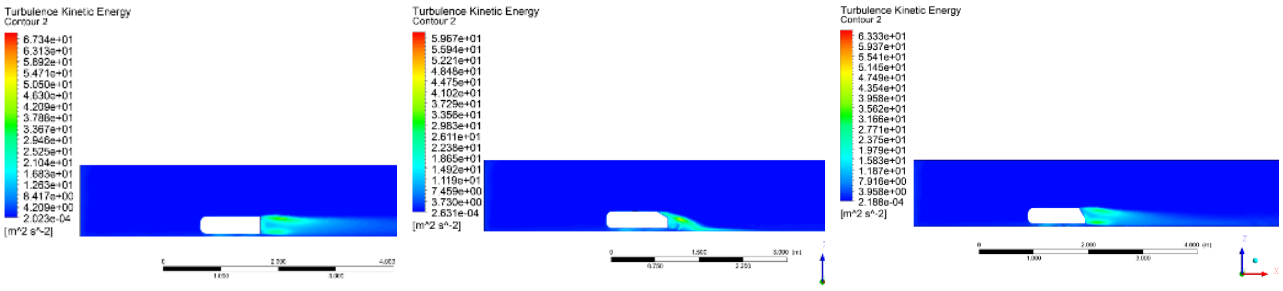
*B. Velocity contour and Turbulence Kinetic Energy plots for the three cases of Ahmed's body (***0**°*,* **30**° *and* **60**° *slant angles):*

The contour colors variation is showing change in velocity. The maximum velocity in 0° slant angle case is 63.23 m/s which is almost same as 60° slant angle case which shows 62.30 m/s but for 30°slant angle, it shows 59.17 m/s. As expected, the highest value occurs top surface of Ahmed's body as the frontal curvature increase to smooth top surface increases velocity as the pressure drag is less on top surface. The nose of the shape has almost zero velocity value as the flow is striking on frontal face and causing pressure drag and has very low velocity at the rear slant side due to formation of vortices. The green area agrees with the velocity value which is 25.363 m/s, while blue area shows low velocity and

 ©2020, IJA-ERA - All Rights Reserved

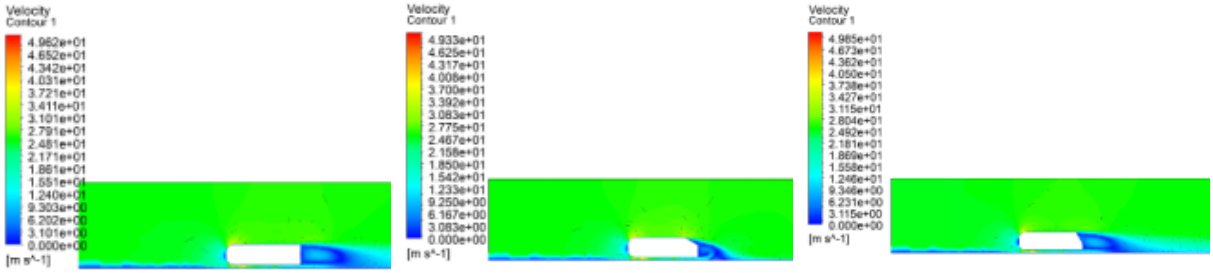
high-pressure region and red area gives high velocity and low-pressure region. Where the pressure is low the velocity increases and where pressure is high the value of velocity decreases. The velocity contours showed that with increase in slant angle up to30°, the flow nearly attached to slant surface in the wake region, and the form drag consequently decreases as the rear slant angle increased from 0° slant angle to 30° slant angle. As the slant angle is increased to 60° the flow is again separated at rear region and flow vortices are significant. The value of lift coefficient and drag coefficient is increased as the slant angle is increased from 0° to 30° as the flow is more attached to surface whereas the slant angle is further increased to 60° the flow separation increases at slant region causing decrease in lift coefficient

When flow becomes turbulent, transport mechanisms like diffusion highly increases beyond the molecular diffusion. For the deviation in slant angles, figure (13) illustrate the thickness of these streams which is determined by turbulent kinetic energy where the flow is separated and in coincidence of the lower horseshoe vortex in 30° slant angle case and wider vortex in 60° slant angle case but still in 30° slant angle case has the wider shape of vortex as flow is more attached to body at slant. The wake region has shown the same results as the velocity contours and that's can be explained as the Kinetic energy is directly proportional to velocity.



a. KTE for 0° slant angle case. b. KTE for 30° slant angle case. c. KTE for 60° slant angle case.

Figure 13: Turbulence Kinetic Energy for the three cases of slant angles.

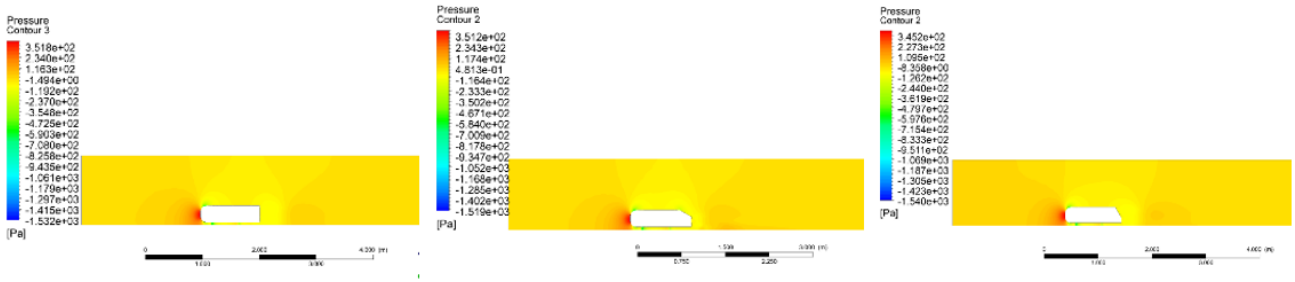


velocity contour for 0° slant angle velocity contour for 30° slant angle velocity contour for 60° slant angle

Figure 14: Velocity contours for the three cases of slant angles.

C. Pressure plots for the three cases of Ahmed's body (0°, 30° and 60° slant angles):

 ©2020, IJA-ERA - All Rights Reserved



a. Pressure contour for 0° slant angle **b. Pressure contour for 30° slant angle** **c. Pressure contour for 60° slant angle**

Figure 15: Pressure contours for the three cases of slant angles.

As the air strikes the frontal area of vehicle, velocity is reduced at frontal face as very high pressure of air acts on the vehicle's front engine's grill which is highlighted in red coloured contour.

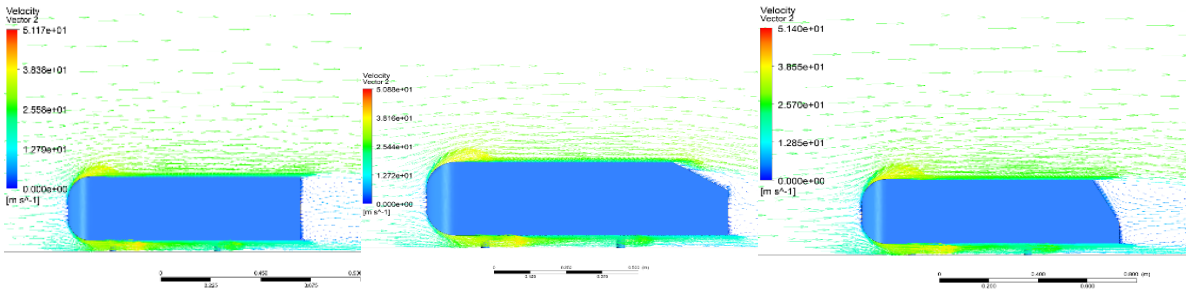
Further, the flow of air gets separated between lower and upper vehicle surfaces. As the air travels over the curved nose surface of Ahmed body, it accelerates with high speed on surface that is forward-facing, causing a drop-in pressure. This low-pressure creates lift over the surface of the roof and the pressure on bottom of vehicle as the air pass on it.

A notch is formed by the back slant owing to separation of flow as the air continue to flow and to the back, leaving a space of low pressure or vacuum where the air won't be able to properly fill as can be seen in figure (15).

D. Velocity vector plots for the three cases of Ahmed's body (0°, 30° and 60° slant angles):

The velocity vector contour shows that the air starts to separate when it strikes with Ahmed's body showing velocity reduces significantly at the frontal region thus it can be expressed that maximum pressure occurs in this region as can be seen in figure (16).

Further, the velocity vectors move along the curves of the shape so that their values increased and then it separated to flow around the body in both sides. The velocity decreases at rear slant region minimum as flow separation occurs that results in to vortices. The vortices change in its width from one case to another and also in its intensity as 0° slant angle case shows the highest intensity and wider vortices than the other cases as shown in figure.



 ©2020, IJA-ERA - All Rights Reserved

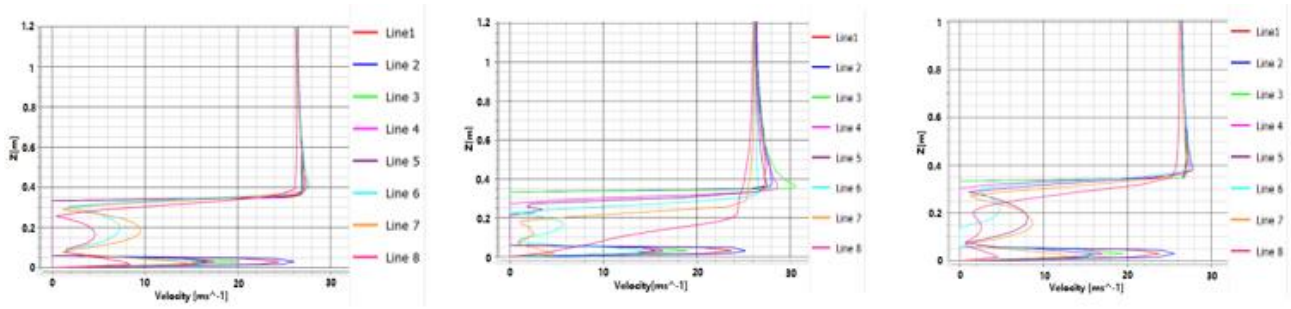
a. Velocity vector for 0°slant angle case b. Velocity vector for 30° slant angle case c. Velocity vector for 60° slant angle case

Figure 16: Velocity vector plots for the three cases of Ahmed's body.

*E. Velocity profiles plots referring to Z axis, for the three cases of Ahmed's body (***0°***,* **30°** *and* **60°** *slant angles):*

The flow over the rear side of Ahmed's body is captured as velocity profiles with the use of lines inserted on rear side of body. The velocity profiles are shown in figure 17, figures (a, b and c) illustrate the flow behavior affected by slant angles in Ahmed's body.

These plots start just before the slant region of Ahmed's body showing the velocity before facing the slant angle and then velocity changes abruptly as it passes the slant region. The change in velocity is nearly same for slant angles of 0° and 60° but for 30° slant angle, it shows significant fluctuation in velocity comparing to other two slant angles and that's prove why the drag coefficient and lift coefficient values in this case are higher than the other two cases .



a. Velocity profiles for 0° slant angle case

b. Velocity profiles for 30° slant angle case

c. Velocity profiles for 60° slant angle case

Figure 17: Velocity profiles for 0º, 30º and 60o slant angles cases.

IV. DISCUSSION

The CFD results obtained the velocity contours, velocity vectors, pressure contours and turbulent kinetic energy contours. It has been known that airflow velocity is inversely proportional to the pressure to keep the algebraic sum of potential energy, kinetic energy, and pressure constant. The high pressure at frontal face of Ahmed's body resulted as the air flow strikes perpendicularly at front face, the air velocity decreases as the result of increase in pressure at frontal region. The air flow continues to flow on upper curved path of Ahmed's body, this curvature provides a way to accelerate the flow on upper surface. Similarly flow passes through bottom surface of Ahmed's body but the flow velocity is less as compared to top surface. This velocity difference of top and bottom surfaces of Ahmed's body causes decrease in pressure at top surface and increase in pressure at lower surface resulting into lift generation. Further, the flow moves toward the rear slant angle in each case and it started to separate forming vortices with high intensity as it's a turbulent flow, but still the vortices strength varied from one case to another as Ahmed's body shape with 0° slant angle shows wider vortices than 30° as flow becomes attached at rear region but when 60° slant angle is studied, it showed that vortex strength is increased again due to flow separation at slant side. The velocity contours results showed that for different slant angles the flow behavior varies at rear region of Ahmed's body. The results illustrated

 ©2020, IJA-ERA - All Rights Reserved

that with increase in slant angle up to30°, the flow is nearly attached to slant surface in the wake region, and the form drag consequently decreases as the rear slant angle is increased from 0° slant angle to 30° slant angle. When the slant angle is further increased to 60° the flow is again separated at rear region and flow vortices are significant. The value of lift coefficient and drag coefficient is increased as the slant angle is increased from 0° to 30° as the flow is more attached to surface whereas the slant angle is further increased to 60° the flow separation increases at slant region causing decrease in lift coefficient. This shows that flow is more attached to Ahmed's body with 30°slant angle configuration. For the difference in slant angles, it is obvious that at the initial section of the rear slant, the pressure increases (evident by the decreasing green area) with decreasing value of rear slant angle from $30^{\circ}$to 0°. This is due to less dramatic change in the direction of airflow, resulting in decreased notch formation (vacuum region or low pressure). This results in more steady pressure between the top and bottom surfaces of the Ahmed body and hence, reduced tendency of lift. The results showed that lift coefficient and drag coefficient increased with increase in rear slant angle from 0° slant angle to 30° slant angle because of the larger pressure difference happened between bottom and top surfaces of the Ahmed's body. On the other hand, the both the drag coefficient and lift coefficient decreased by increasing rear slant angle is increased from 30° slant angle to 60° slant angle as the flow is again detached at rear slant region causing flow separation resulting in loss of lift. Shape effects on drag are presented in figure (18).

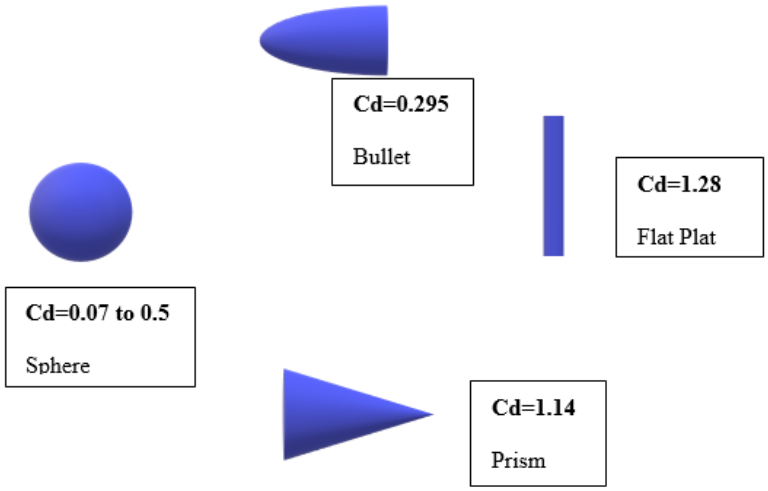


Figure 18: Shape effects on Drag

In figure (18), we can easily see that at same frontal area, then, C_d is depends upon only shape

From the observations, we found that at 30^0 slant angle , value of C_d is very high as comparison to other two slant angles like 60^0 and 0^0.Hence , 30^0 is not a suitable choice as it will consume more fuel and not so efficient.

Table 4: Lift and Drag coefficient convergence for flow past Ahmed's body cases.

Slant Angles	C_d	C_L
$\theta=0^{\circ}$	0.15106	0.57380

 ©2020, IJA-ERA - All Rights Reserved

$\theta=30^{\circ}$	0.16146	0.59823
$\theta=60^{\circ}$	0.15010	0.57765

We can see drag and lift is maximum at 30-degree slant angles as comparison to other slant angles case. This is occurring due to couple of vortices developed, and the flow leave the body at the lower edge of the slant. This leads to very strong under pressures at that position, so that the coefficient of drag reaches its maximum in this position at angle $=30^{0}$ The difference between lift coefficient and drag coefficient values of Ahmed's body configurations of 0° slant angle to 30° slant angle. Whereas for the 60° slant angle configuration the lift coefficient is increased to 7 percent and drag coefficient is increased to 4 percent as compared to 0° slant angle configuration.

V. CONCLUSION

The CFD analysis of Ahmed's body with slant angles of 0°, 30° and 60° is studied using ANSYS FLUENT with same fluid flow conditions. The Design Modeler tool of ANSYS is used for CAD model preparation of Ahmed's body different cases. As the Reynold's number lied in turbulent k-epsilon two equation turbulent model is used with transient flow analysis. The results of velocity contour, pressure contour, Kinetic Turbulent Energy, velocity vector and velocity profile plots are presented for analysis of aerodynamic behavior of Ahmed's body. The results showed that by varying slant angle from $0^{\circ}$to $30^{\circ}$the flow is attached at rear slant region and flow separation is reduced but lift coefficient and drag coefficient values are increased. Whereas increasing the slant angle to 60°, flow again detached the flow at rear region causing flow separation but there is not much difference in lift and drag coefficients as comparing to 0°slant angle configuration. The following conclusions are drawn from results:

1. The minimum drag coefficient was 0.15010 is obtained in Ahmed's Body with 60°rear slant angle.
2. The minimum lift coefficient was 0.5738 is obtained in Ahmed's Body with 0° degree rear slant angle.
3. The lift coefficient and drag coefficient increased by 4%, and 7% respectively as increasing the slant angle from $0^{\circ}$to 30°. The boundary layer separation occurs on the rear slant caused increase in drag coefficient. However, Lift coefficient increased with increasing in rear slant angle due to larger pressure difference formed.
4. The flow separation vortices at rear region of wake area decreased with increase in slant angle up to 30°. This happened at rear end as a result of pressure difference, the flow on top and bottom of Ahmed's body joined at trailing region thus reducing drag causing vortices. As there are small vortices at back of Ahmed's body so less dissipation of kinetic energy takes place. The vehicle engine will require less power for propelling car in forward direction.
5. Wake is a region of recirculated flow which induces low pressure behind the car which is the main cause of the aerodynamic drag of car.

VI. FUTURE SCOPE OF STUDIES

Here, we discussed about the effect of slant angles on the coefficient of drag & coefficient of lift. It is fully optimized but somehow, we can say other factors can also affect C_d drastically. Future researchers may work on the shape of the Ahmed body and other slant angles to see some more results at which slant angle, drag and lift is more.

 ©2020, IJA-ERA - All Rights Reserved

Conflict of interest: The authors declare that they have no conflict of interest.

Ethical statement: The authors declare that they have followed ethical responsibilities

REFERENCES

[1] Zulfaa Mohamed-Kassim, Antonio Filippone, Fuel savings on a heavy vehicle via aerodynamic drag reduction,Transportation Research Part D: Transport and Environment,Volume 15, Issue 5,2010,Pages 275-284,ISSN 1361-9209,doi.org/10.1016/j.trd.2010.02.010

[2] Blazek, J. (2015). Computational Fluid Dynamics: Principles and Applications: Third Edition.In Computational Fluid Dynamics:Principles and Applications: Third Edition.doi.org/10.1016/C2013-0-19038-1

[3] Costa, B. (2015). Studying the Airflow Over a Car Using an Ahmed Body. In Comsol (p. Page: 23.). https://www.researchgate.net/publication/325451944_CFD_Study_of_Flow_over_a_Simplified_Car_Ahmed_Body_Using_Different_Turbulence_Models

[4] Ahmed, S. R., Ramm, G., & Faltin, G. (1984). Some salient features of the time-averaged ground vehicle wake. SAE Technical Papers, Pages: 71_77. doi.org/10.4271/840300

[5] Bayraktar, I., Landman, D., & Baysal, O. (2001). Experimental and computational investigation of Ahmed body for ground vehicle aerodynamics. SAE Technical Papers,Page:2_3. doi.org/10.4271/2001-01-2742

[6] Lienhart, H., & Becker, S. (2003). Flow and turbulence structure in the wake of a simplified car model. SAE Technical Papers, Page: 2_3

[7] Bello-Millán, F. J., Mäkelä, T., Parras, L., del Pino, C., & Ferrera, C. (2016). Experimental study on Ahmed's body drag coefficient for different yaw angles. Journal of Wind Engineering and Industrial Aerodynamics, 157, 140–144. doi:10.1016/j.jweia.2016.08.005

[8] Hinterberger, C., M. García-Villalba, W. Rodi (2004) Large eddy simulation of flow around the Ahmed body. In: McCallen R, Browand F, Ross J (Eds.) Lecture notes in applied and computational mechanics / The aerodynamics of heavy vehicles: trucks, busses, and trains, Springer Verlag

[9] Guilmineau, E. (2008). Computational study of flow around a simplified car body. Journal of Wind Engineering and Industrial Aerodynamics,96(6–7),1207_1217. doi.org/10.1016/j.jweia.2007.06.041

[10] Maheshkumar Maruti Patil, Mariappan D. Nadar, Suhas A. Uthale(2016) To investigate Heat Loss of a Fluid flowing through a Pipeline for Turbulent Flow. International Journal of Advanced Engineering Research and Applications (IJAERA)

 ©2020, IJA-ERA - All Rights Reserved

View publication stats

arXiv:2009.02719v1 [math.CV] 6 Sep 2020

ON GEOMETRICAL PROPERTIES OF CERTAIN ANALYTIC FUNCTIONS

S. SIVAPRASAD KUMAR AND KAMALJEET GANGANIA

Abstract. We introduce the class of analytic functions

$$\mathcal{F}(\psi) := \left\{ f \in \mathcal{A} : \left(\frac{zf'(z)}{f(z)} - 1 \right) \prec \psi(z),\ \psi(0) = 0 \right\},$$

where ψ is univalent and establish the growth theorem with some geometric conditions on ψ and obtain the Koebe domain with some related sharp inequalities. Note that functions in this class may not be univalent. As an application, we obtain the growth theorem for the complete range of α and β for the functions in the classes $\mathcal{BS}(\alpha) := \{f \in \mathcal{A} : (zf'(z)/f(z)) - 1 \prec z/(1-\alpha z^2),\ \alpha \in [0,1)\}$ and $\mathcal{S}_{cs}(\beta) := \{f \in \mathcal{A} : (zf'(z)/f(z)) - 1 \prec z/((1-z)(1+\beta z)),\ \beta \in [0,1)\}$, respectively which improves the earlier known bounds. The sharp Bohr-radii for the classes $S(\mathcal{BS}(\alpha))$ and $\mathcal{BS}(\alpha)$ are also obtained. A few examples as well as certain newly defined classes on the basis of geometry are also discussed.

2010 AMS Subject Classification. Primary 30C80, Secondary 30C45.
Keywords and Phrases. Subordination, Bohr-Radius, Majorization, Distortion theorem.

1. Introduction

Let $\mathcal{A}$ denotes the class of analytic functions of the form $f(z) = z + \sum_{k=2}^{\infty} a_k z^k$ in the open unit disk $\Delta := \{z : |z| < 1\}$. Let $f(z) = w$ and Γ_w be the image of Γ_z (the circle $C_r : z = re^{i\theta}$) under the function f in $\mathcal{A}$. The curve Γ_w is said to be starlike with respect to $w_0 = 0$ if $\arg(w - w_0)$ is a non-decreasing function of θ, that is,

$$\frac{d}{d\theta} \arg(w - w_0) \geq 0, \quad \theta \in [0, 2\pi],$$

which is equivalent to

$$\frac{d}{d\theta} \arg(w - w_0) = \Re \left(\frac{zf'(z)}{f(z)} \right) \geq 0. \tag{1.1}$$

If the inequality (1.1) holds for each circle $|z| = r < 1$, then it characterizes a special class $\mathcal{S}^*$, the class of starlike functions in Δ. It is obvious from (1) that for each $0 < r < 1$, the curve Γ_w is not allowed to have a loop which ensure the univalency of the function. But if for some $0 \neq z \in \Delta$, $\Re(zf'(z)/f(z)) < 0$, then f is not starlike with respect to 0, or equivalently we can say that the image curve $\Gamma_w : f(|z| = r)$ is definitely not starlike, but still it may or may not be univalent. From (1.1), we also see the importance of the Caratheódory functions by writing (1.1) in terms of

The work of the second author is supported by University Grant Commission, New-Delhi, India under UGC-Ref. No.:1051/(CSIR-UGC NET JUNE 2017).

subordination as:

$$\frac{zf'(z)}{f(z)} \prec \frac{1+z}{1-z} \quad (z \in \Delta), \tag{1.2}$$

where the symbol $\prec$ stands for the usual subordination. In 1992, Ma-Minda [8] generalized (1.2) by unifying all the subclasses of starlike functions as follows:

$$\mathcal{S}^*(\Psi) := \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} \prec \Psi(z) \right\}, \tag{1.3}$$

where Ψ has positive real part, $\Psi(\Delta)$ symmetric about the real axis with $\Psi'(0) > 0$ and $\Psi(0) = 1$. For some special classes, refer [4, 5, 14] and the references there in.

In view of the above, let us now consider the analytic univalent function ψ in Δ such that $\psi(0) = 0$, $\psi(\Delta)$ is starlike with respect to 0 and introduce the following class of analytic functions:

$$\mathcal{F}(\psi) := \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} - 1 \prec \psi(z),\ \psi(0) = 0 \right\}. \tag{1.4}$$

Note that when $1 + \psi(z) \not\prec (1+z)/(1-z)$, then the functions in the class $\mathcal{F}(\psi)$ may not be univalent in Δ which also implies $\mathcal{F}(\psi) \not\subseteq \mathcal{S}^*$. Thus in case, when the function $1 + \psi := \Psi$ has positive real part, $\Psi(\Delta)$ symmetric about the real axis with $\Psi'(0) > 0$, then $\mathcal{F}(\psi)$ reduces to the class $\mathcal{S}^*(\Psi)$. The functions in the class defined in (1.3) are univalent which help a lot in studying the geometrical properties of the functions, for example, Growth and Distortion theorems etc. But this may not be quite easy to study the analogous results in the class $\mathcal{F}(\psi)$. In this direction, recently, Kargar et al. [6] considered the following class, the first of it's kind:

$$\mathcal{BS}(\alpha) := \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} - 1 \prec \frac{z}{1-\alpha z^2},\ \alpha \in [0,1) \right\}, \tag{1.5}$$

where $z/(1-\alpha z^2) =: \psi(z)$ (Booth lemniscate function [11] and [12]) is an analytic univalent function and symmetric with respect to the real and imaginary axes. Note that the function $(1 + z/(1-\alpha z^2))$ assumes negative values for $\alpha \neq 0$, therefore functions in this class may not be univalent. For f belonging to $\mathcal{BS}(\alpha)$, using the vertical strip domain $\{w \in \mathbb{C} : \mu < \Re w < \nu, \text{ where } \mu < 1 < \nu\}$, Kargar et al. [6] proved that $|f(z)/z|$ is bounded and obtained the coefficient estimates when $0 \leq \alpha \leq 3 - 2\sqrt{2}$ along with Fekete-Szegö inequality for the associated $k-th$ root transformation. In 2018, Najmadi et al. [10] obtained the bounds for the quantities $\Re f(z)$, $|f(z)|$ and $|f'(z)|$ when $0 \leq \alpha \leq 3-2\sqrt{2}$. Recently, Kargar et al. [7] obtained the best dominant of the subordination $f(z)/z \prec F(z)$ for the range $0 < \alpha \leq 3-2\sqrt{2}$ using the convolution technique, where $F(z) = (1 + z\sqrt{\alpha})/(1 - z\sqrt{\alpha})^{\frac{1}{2\sqrt{\alpha}}}$. Cho et al. [3] dealt with the first order differential subordination implications and also solved the various sharp radius problems pertaining to the class $\mathcal{BS}(\alpha)$.

In 2019, Masih et al. [9] considered the following class with $\beta \in [0, 1/2]$:

$$\mathcal{S}_{cs}(\beta) := \left\{ f \in \mathcal{A} : \left(\frac{zf'(z)}{f(z)} - 1 \right) \prec \frac{z}{(1-z)(1+\beta z)},\ \beta \in [0,1) \right\}. \tag{1.6}$$

They proved the growth theorem and also obtained the sharp estimates for the logarithmic coefficients but only for the range $\beta \in [0, 1/2]$. Note that for $\beta \in [0, 1/2]$, $\mathcal{S}_{cs}(\beta)$ is a Ma-Minda subclass, but for the other range, functions in this class may not be univalent.

In this paper, we establish the sharp growth theorem for the class $\mathcal{F}(\psi)$ with certain geometric conditions on ψ and obtain covering theorem. Further provide some examples including newly defined classes are also discussed. As an application, we obtain growth theorem for the complete range of α and β for the functions in the classes $\mathcal{BS}(\alpha)$ and $\mathcal{S}_{cs}(\beta)$, respectively that improves the earlier known bounds. Finally, the sharp Bohr-radii for the classes $S(\mathcal{BS}(\alpha))$ and $\mathcal{BS}(\alpha)$ are obtained. For some classes, we study the geometrical behavior of an analytic function of the form $f(z)/z$ which arises frequently while working with the class $\mathcal{S}^*(\Psi)$ and play an important role, for example, in obtaining the bounds for $\Re(f(z)/z)$ and $\arg(f(z)/z)$. The geometric properties and coefficients estimation for the class $\mathcal{F}(\psi)$ are still open.

2. Main Results

Let $\mathcal{F}(\psi)$ be the class as defined in (1.4). Now we begin with the following:

Theorem 2.1 (Growth Theorem)**.** *If* $\max_{|z|=r} \Re\psi(z) = \psi(r)$ *and* $\min_{|z|=r} \Re\psi(z) = \psi(-r)$*. Then* $f \in \mathcal{F}(\psi)$ *satisfies the sharp inequalities:*

$$r \exp\left(\int_0^r \frac{\psi(-t)}{t} dt\right) \leq |f(z)| \leq r \exp\left(\int_0^r \frac{\psi(t)}{t} dt\right), \quad (|z| = r). \tag{2.1}$$

Proof. Let $f \in \mathcal{F}(\psi)$. For $z = re^{i\theta}$, we have

$$\phi(-r) \leq \Re\psi(re^{i\theta}) \leq \phi(r). \tag{2.2}$$

Let $\Phi(z) = \psi(\omega(z))$, where ω is a Schwarz function. Then from (1.3), we have

$$\log \frac{f(z)}{z} = \int_0^z \frac{\Phi(\zeta)}{\zeta} d\zeta.$$

Now by taking $\zeta = te^{i\beta}$ so that $d\zeta = e^{i\beta} dt$, where β is fixed but arbitrary and $z = re^{i\beta}$, we have

$$\log \frac{f(z)}{z} = \int_0^r \frac{\Phi(te^{i\beta})}{t} dt. \tag{2.3}$$

From the Maximum-minimum modulus principle, we find that Φ also satisfies the inequality (2.2). Therefore, without loss of generality, we may replace Φ by ψ and β by θ in (2.3). Then by equating real parts on either side of (2.3), we have

$$\log \left|\frac{f(z)}{z}\right| = \int_0^r \frac{\Re\Phi(te^{i\theta})}{t} dt \tag{2.4}$$

and now using the inequalities (2.2) in (2.4), we obtain

$$\int_0^r \frac{\psi(-t)}{t} dt \leq \log \left|\frac{f(z)}{z}\right| \leq \int_0^r \frac{\psi(t)}{t} dt,$$

and (2.1) follows. The result is sharp for the function

$$f_0(z) = z \exp \int_0^z \frac{\psi(t)}{t} dt. \tag{2.5}$$

This completes the proof.

Remark 2.1. In the above theorem, we chose $\max_{|z|=r} \Re\psi(z) = \psi(r)$ and $\min_{|z|=r} \Re\psi(z) = \psi(-r)$ for computational convenience. However, these conditions may change according to the choice of ψ in that case, appropriately these may be replaced.

Remark 2.2. If $1+\psi$ is a Carathéodory univalent function, then Theorem 2.1 reduces to the result [8, Corollary 1, p. 161] and moreover, letting r tends to 1 in Theorem 2.1, we obtain the covering theorem (Koebe-radius) for the class $\mathcal{F}(\psi)$.

Corollary 2.2 (Covering Theorem)**.** *If $f \in \mathcal{F}(\psi)$ and f_0 as defined in* (2.5)*, then either f is a rotation of f_0 or*

$$\{w \in \Delta : |w| \leq -f_0(-1)\} \subset f(\Delta),$$

where $-f_0(-1) = \lim_{r\to 1}(-f_0(-r))$.

Let $L(f, r)$ denotes the length of the boundary curve $f(|z| = r)$. Note that for $z = re^{i\theta}$, we have $L(f, r) := \int_0^{2\pi} |zf'(z)|d\theta$. Now we obtain the following result:

Corollary 2.3. *Assume that $\max_{|z|=r} |\psi(z)| = \psi(r)$ and also ψ satisfies the conditions of Theorem 2.1. Let $M(r) = \exp\left(\int_0^r \frac{\psi(t)}{t}dt\right)$. If $f \in \mathcal{F}(\psi)$, then for $|z| = r$, we have*

$$\Re\frac{f(z)}{z} \leq M(r), \;\; |f'(z)| \leq (1 + \psi(r))M(r) \;\; \textit{and} \;\; L(f, r) \leq 2\pi r(1 + \psi(r))M(r).$$

Let

$$\psi(z) = \begin{cases} \beta z/(1 + \alpha z), & \beta > 0,\ 0 < \alpha < 1\ ; \\ \eta z, & \eta > 0. \end{cases}$$

Then the above two choices of ψ are clearly convex univalent and $\psi(\Delta)$ are symmetric about real axis as $\overline{\psi(z)} = \psi(\bar{z})$. It is further evident that $1 + \psi(z) \not\prec (1 + z)/(1 - z)$ except for the second choice of ψ when $0 < \eta \leq 1$. We now obtain the following sharp result from Theorem 2.1:

Example 2.4. *Let $f \in \mathcal{F}(\beta z/(1 + \alpha z))$, where $\beta > 0$ and $0 < \alpha < 1$ and $|z| = r$. Then*

$$r(1 - \alpha r)^{\frac{\beta}{\alpha}} \leq |f(z)| \leq r(1 + \alpha r)^{\frac{\beta}{\alpha}},$$

which implies:

$$\left\{w : |w| \leq (1 - \alpha)^{\frac{\beta}{\alpha}}\right\} \subset f(\Delta), \;\; |f'(z)| \leq \left(1 + \frac{\beta r}{1 + \alpha r}\right)(1+\alpha r)^{\frac{\beta}{\alpha}} \;\; \textit{and} \;\; \Re\frac{f(z)}{z} \leq (1+\alpha r)^{\frac{\beta}{\alpha}}.$$

Example 2.5. *Let $f \in \mathcal{F}(\eta z)$, where $\eta > 0$ and $|z| = r$. Then*

$$r\exp(-\eta r) \leq |f(z)| \leq r\exp(\eta r),$$

which implies:

$$\{w : |w| \leq \exp(-\eta)\} \subset f(\Delta), \;\; |f'(z)| \leq (1 + \eta r)\exp(\eta r) \;\; \textit{and} \;\; \Re\frac{f(z)}{z} \leq \exp(\eta r).$$

From the above examples, it is clear that $f \in \mathcal{F}(\psi)$ if and only if

$$\frac{zf'(z)}{f(z)} \in \begin{cases} \Omega_1, & \text{when } \psi(z) = \beta z/(1 + \alpha z); \\ \Omega_2, & \text{when } \psi(z) = \eta z, \end{cases}$$

where $\Omega_1 = \{w \in \mathbb{C} : |w - 1| < |\beta - \alpha(w - 1)|\}$ and $\Omega_2 = \{w \in \mathbb{C} : |w - 1| < \eta\}$, respectively for $z \in \Delta$.

3. Some Applications and Further results

3.1. **On Booth-Lemniscate.** Let $\mathcal{BS}(\alpha)$ be the class as defined in (1.5).

Theorem 3.1. *Let $0 < \alpha < 1$ and $f \in \mathcal{BS}(\alpha)$, then for $|z| = r$*

$$-\hat{f}(-r) \leq |f(z)| \leq \hat{f}(r), \tag{3.1}$$

where

$$\hat{f}(z) = z\left(\frac{1+z\sqrt{\alpha}}{1-z\sqrt{\alpha}}\right)^{\frac{1}{2\sqrt{\alpha}}}. \tag{3.2}$$

The result is sharp.

Proof. Let $\psi(z) := z/(1-\alpha z^2)$ and $f \in \mathcal{BS}(\alpha) := \mathcal{F}(\psi)$. For $z = re^{i\theta}$, we have

$$-\frac{r}{1-\alpha r^2} \leq \Re\psi(re^{i\theta}) \leq \frac{r}{1-\alpha r^2}$$

and

$$-\int_0^r \frac{1}{1-\alpha t^2}dt \leq \log\left|\frac{f(z)}{z}\right| \leq \int_0^r \frac{1}{1-\alpha t^2}dt,$$

where

$$\int_0^r \frac{1}{1-\alpha t^2}dt = \frac{1}{2\sqrt{\alpha}}\log\frac{1+\sqrt{\alpha}r}{1-\sqrt{\alpha}r}.$$

Hence, the result follows from Theorem 1.4.

Remark 3.1. Theorem 3.1 improves the upper bound of $\Re f(z)$ and bounds of $|f(z)|$, obtained in [10, Theorem 2, p. 116] and [10, Theorem 3, p. 116] respectively.

We now extend [7, Theorem 2.6, p. 1238] for the complete range of α using Theorem 3.1:

Corollary 3.2. *Let $f \in \mathcal{BS}(\alpha)$, $\alpha \in (0,1)$ and $|z| = r$, then*

$$\Re\frac{f(z)}{z} \leq \left(\frac{1+r\sqrt{\alpha}}{1-r\sqrt{\alpha}}\right)^{\frac{1}{2\sqrt{\alpha}}} \quad \text{and} \quad |f'(z)| \leq \left(1+\frac{r}{1-\alpha r^2}\right)\left(\frac{1+r\sqrt{\alpha}}{1-r\sqrt{\alpha}}\right)^{\frac{1}{2\sqrt{\alpha}}}.$$

The result is sharp for the function $\hat{f}$ given in (3.2).

Corollary 3.3. *Let $\alpha \in (0,1)$ be fixed. Then $f \in \mathcal{BS}(\alpha)$ satisfies the inequality*

$$L(f,r) \leq 2\pi r\left(1+\frac{r}{1-\alpha r^2}\right)\left(\frac{1+r\sqrt{\alpha}}{1-r\sqrt{\alpha}}\right)^{\frac{1}{2\sqrt{\alpha}}}, \quad (|z| = r).$$

Corollary 3.4 (Koebe-radius)**.** *Let $0 < \alpha < 1$ and $\hat{f}$ as given in* (3.2). *If $f \in \mathcal{BS}(\alpha)$, then either f is a rotation of $\hat{f}$ or*

$$\{w \in \mathbb{C} : |w| \leq -\hat{f}(-1)\} \subset f(\Delta).$$

Proof. The proof follows by letting r tends to 1 in the inequality $-\hat{f}(-r) \leq |f(z)|$, given in (3.1).

Theorem 3.5. *Let $\alpha \in (0, 2-\sqrt{3}]$ be fixed. Then $f \in \mathcal{BS}(\alpha)$ satisfies the sharp inequality*

$$\left|\arg\frac{f(z)}{z}\right| \leq \max_{|z|=r} \arg\left(\frac{1+z\sqrt{\alpha}}{1-z\sqrt{\alpha}}\right)^{\frac{1}{2\sqrt{\alpha}}}.$$

Proof. From [7, Theorem 2.5, p. 1238], we have $f(z)/z \prec \hat{f}(z)/z$ for $0 < \alpha \leq 2-\sqrt{3}$, where $\hat{f}$ is defined in (3.2). Since the function $\hat{f}(z)/z$ is convex and symmetric about the real axis in Δ, therefore we easily see that

$$\left(\frac{1-\sqrt{\alpha}}{1+\sqrt{\alpha}}\right)^{\frac{1}{2\sqrt{\alpha}}} > 0.$$

Thus $\hat{f}(z)/z$ is a Caratheódory function and the result follows.

For our next result, we need the following definition and a related class:

Definition 3.6. Let $f(z) = \sum_{k=0}^{\infty} a_k z^k$ and $g(z) = \sum_{k=0}^{\infty} b_k z^k$ are analytic in Δ and $f(\Delta) = \Omega$. Consider a class of analytic functions $S(f) := \{g : g \prec f\}$ or equivalently $S(\Omega) := \{g : g(z) \in \Omega\}$. Then the class $S(f)$ is said to satisfy Bohr-phenomenon, if there exists a constant $r_0 \in (0, 1]$ such that the inequality $\sum_{k=1}^{\infty} |b_k| r^k \leq d(f(0), \partial\Omega)$ holds for all $|z| = r \leq r_0$, where $d(f(0), \partial\Omega)$ denotes the Euclidean distance between $f(0)$ and the boundary of $\Omega = f(\Delta)$. The largest such r_0 for which the inequality holds, is called the Bohr-radius.

See the articles [1, 2] and the references therein for more. Let us now introduce the following class:

$$S(\mathcal{BS}(\alpha)) := \left\{g : g \prec f,\ g(z) = \sum_{k=1}^{\infty} b_k z^k \text{ and } f \in \mathcal{BS}(\alpha)\right\}.$$

Theorem 3.7 (Booth-Bohr-radius)**.** *The class $S(\mathcal{BS}(\alpha))$ satisfies Bohr-phenomenon in $|z| \leq r(\alpha)$, where $r(\alpha)$ is the unique positive root of the equation*

$$r\left(\frac{1+r\sqrt{\alpha}}{1-r\sqrt{\alpha}}\right)^{\frac{1}{2\sqrt{\alpha}}} - \left(\frac{1-\sqrt{\alpha}}{1+\sqrt{\alpha}}\right)^{\frac{1}{2\sqrt{\alpha}}} = 0, \tag{3.3}$$

whenever $0 < \alpha \leq 3 - 2\sqrt{2}$. The result is sharp for the function $\hat{f}$ given in (3.2)*.*

Proof. Since $g \in S(\mathcal{BS}(\alpha))$, we have $g \prec f$ for a fixed $f \in \mathcal{BS}(\alpha)$. From Corollary 3.4, we obtain the Koebe-radius $r_* = -\hat{f}(-1)$ such that $r_* \leq d(0, \partial\Omega) = |f(z)|$ for $|z| = 1$. Also using [7, Theorem 2.5, p. 1238], we have

$$\frac{f(z)}{z} \prec \frac{\hat{f}(z)}{z}. \tag{3.4}$$

Recall the result [2, Lemma 1, p.1090], which reads as: let f and g be analytic in Δ with $g \prec f$, where $f(z) = \sum_{n=0}^{\infty} a_n z^n$ and $g(z) = \sum_{k=0}^{\infty} b_k z^k$. Then $\sum_{k=0}^{\infty} |b_k| r^k \leq \sum_{n=0}^{\infty} |a_n| r^n$ for $|z| = r \leq 1/3$. Now using the result for $g \prec f$ and (3.4), we have

$$\sum_{k=1}^{\infty} |b_k| r^k \leq r + \sum_{n=2}^{\infty} |a_n| r^n \leq \hat{f}(r) \quad \text{for } |z| = r \leq 1/3.$$

Finally, to establish the inequality $\sum_{k=1}^{\infty} |b_k| r^k \leq d(f(0), \partial\Omega)$, it is enough to show $\hat{f}(r) \leq r_*$. But this holds whenever $r \leq r(\alpha)$, where $r(\alpha)$ is the least positive root of the equation $\hat{f}(r) = r_*$. Now let $T(r) := \hat{f}(r) - r_*$, then

$$T'(r) = \left(\frac{1+r\sqrt{\alpha}}{1-r\sqrt{\alpha}}\right)^{\frac{1}{2\sqrt{\alpha}}} + r\left(\frac{1+r\sqrt{\alpha}}{1-r\sqrt{\alpha}}\right)^{\frac{1}{2\sqrt{\alpha}}-1} \frac{1}{(1-r\sqrt{\alpha})^2}.$$

Since $(1 + r\sqrt{\alpha})/(1 - r\sqrt{\alpha}) > 0$, therefore $T'(r) > 0$ and so T is an increasing function of r. Also $T(0) < 0$ and $T(1) > 0$. Thus the existence of the root $r(\alpha)$

is ensured by the Intermediate Value theorem for the continuous functions. By a computation, it can easily be seen that $r(\alpha) < 1/3$ and hence the result.

Corollary 3.8. *Let $0 < \alpha \leq 3-2\sqrt{2}$. The Bohr-radius for the class $\mathcal{BS}(\alpha)$ is $r(\alpha)$, where $r(\alpha)$ is the unique positive root of the Eq.* (3.3).

3.2. **On Cissoid of Diocles.** Let us consider

$$S_\beta(z) = \frac{z}{(1-z)(1+\beta z)} = \frac{1}{1+\beta}\left(\frac{1}{1-z} + \frac{1}{1+\beta z}\right) = \sum_{n=1}^{\infty} \frac{1-(-\beta)^n}{1+\beta} z^n,$$

where $\beta \in [0,1)$. Clearly, it is analytic, symmetric about the real-axis and maps the unit disk Δ onto the domain bounded by *Cissoid of Diocles*:

$$CS(\beta) := \left\{ w = u + iv \in \mathbb{C} : \left(u - \frac{1}{2(\beta-1)}\right)(u^2+v^2) + \frac{2\beta}{(1+\beta)^2(\beta-1)} v^2 = 0 \right\}.$$

Let us now consider the class $\mathcal{S}_{cs}(\beta)$ as defined in (1.6). Masih et al. [9] considered this class with $\beta \in [0, 1/2]$ since $\Re(1+z/((1-z)(1+\beta z)) \geq (2\beta-1)/(2(\beta-1)) \geq 0$. Clearly, $\mathcal{S}_{cs}(\beta) = \mathcal{F}(S_\beta(z))$ for $\beta \in [0,1)$ and we have the following result:

Theorem 3.9. *Let $f \in \mathcal{S}_{cs}(\beta)$ and $\beta \in [0,1)$. Then*

$$-\tilde{f}(-r) \leq |f(z)| \leq \tilde{f}(r),$$

where

$$\tilde{f}(z) = z\left(\frac{1+\beta z}{1-z}\right)^{\frac{1}{1+\beta}}. \tag{3.5}$$

Proof. Let $\psi(z) := z/((1-z)(1+\beta z))$ and $f \in \mathcal{S}^*_{cs}(\beta) := \mathcal{F}(\psi)$. Following the proof of [9, Theorem 3.1, p. 5], it is easy to see that for $z = re^{i\theta}$, where $\theta \in [0, 2\pi]$, we have

$$\min_{|z|=r} \Re\psi(z) = \frac{-r + (\beta-1)r^2 + \beta r^3}{(1+r)^2(1-\beta r)^2} = \psi(-r)$$

and

$$\begin{aligned}\max_{|z|=r} \Re\psi(z) &= \lim_{\theta\to 0} \frac{-r^2 + \beta r^2 - \beta r^3\cos\theta + r\cos\theta}{(1+r^2-2r\cos\theta)(1+\beta^2 r^2 + 2\beta r\cos\theta)} \\ &\leq \frac{\beta-1}{2(1+\beta)^2} = \max_{|z|=1} \Re\psi(z).\end{aligned}$$

Thus, we have $\psi(-r) \leq \Re\psi(z) \leq \psi(r)$ for $r \neq 1$ and $1/(2(\beta-1)) = \psi(-1) \leq \Re\psi(z) \leq (\beta-1)/(2(\beta+1)^2)$ for $r = 1$. Also note that

$$\tilde{f}(z) = z\exp\int_0^z \frac{\psi(t)}{t}dt = z\left(\frac{1+\beta z}{1-z}\right)^{\frac{1}{1+\beta}}.$$

Now the result follows from Theorem 1.4.

Remark 3.2. Let $\tilde{F}(z) = \tilde{f}(z)/z$ and $|z| = 1$, where $\tilde{f}$ is as defined in Theorem 3.9. A calculation show that

$$1 + \frac{\tilde{F}''(z)}{\tilde{F}'(z)} = 1 + \frac{-\beta z}{(1+\beta z)(1-z)} + \frac{2z}{1-z},$$

which implies that

$$\Re\left(1 + \frac{\tilde{F}''(z)}{\tilde{F}'(z)}\right) \geq \beta\Re\left(\frac{-z}{(1+\beta z)(1-z)}\right).$$

Since

$$\Re\left(\frac{-z}{(1+\beta z)(1-z)}\right) = \frac{1-\beta}{2(1-\beta^2+2\beta\cos\theta)} =: g(\theta),$$

and a simple calculation shows that g attains its minimum at $\theta = 0$. Therefore, we have

$$1 + \frac{\tilde{F}''(z)}{\tilde{F}'(z)} \geq \frac{\beta(1-\beta)}{2(1+\beta)^2} \geq 0.$$

Hence $\tilde{F}$ is convex univalent in Δ.

Observe that the function $S_\beta(z)$ is not convex when $\beta \neq 0$ and the result, $f(z)/z \prec \tilde{F}(z)$ similar to theorem 3.14 is still open for $f \in \mathcal{S}_{cs}(\beta)$. By letting r tends to 1 in the above Theorem 3.9, we obtain:

Corollary 3.10 (Koebe-radius)**.** *Let $\tilde{f}$ as given in* (3.5)*. If $f \in \mathcal{S}_{cs}(\beta)$, then either f is a rotation of $\tilde{f}$ or*

$$\left\{w \in \mathbb{C} : |w| \leq -\tilde{f}(-1) = \left(\frac{1-\beta}{2}\right)^{1/(1+\beta)}\right\} \subset f(\Delta).$$

Remark 3.3. We improved the result [9, Corollary 4.3.1, p. 8] in Theorem 3.9 and Corollary 3.10 by extending the range of β.

3.3. **Modified Koebe function:** The Koebe function $k(z) = z/(1-z)^2$ has a pole at $z = 1$ and maps unit disk onto the domain $\mathbb{C} - (-\infty, 1/4]$, which is a slit domain. We now introduced the modified Koebe function:

$$K(z) := \frac{z}{(1+\eta z)^2}, \quad 0 \leq \eta < 1, \tag{3.6}$$

which is bounded in Δ and symmetric about the real-axis. It is interesting to observe the geometry of the domain $K(\Delta)$, which assumes different shapes for different choices of η such as a convex or a Bean or a Cardioid shaped domain. Especially when η tends to 1, we see that one of the rotation of the image domain $K(\Delta)$ will converge to $k(\Delta)$ and thereby justifying the name of $K(z)$. Since $k(z) = (u^2(z) - 1)/4$, where $u(z) = (1+z)/(1-z)$, in a similar fashion, we can write

$$K(z) = \frac{1}{4\eta}(1 - v^2(z)),$$

where $v(z) = (1-\eta z)/(1+\eta z)$ and $\eta \neq 0$.

Lemma 3.1. *The function $K(z)$ as defined in* (3.6) *is convex for $0 \leq \eta \leq 2 - \sqrt{3}$.*

Proof. Let $K(z) = z/(1+\eta z)^2$. When $\eta = 0$, $K(z)$ is the identity function and hence is convex. So let us consider $0 < \eta < 1$. By a computation, we obtain that

$$1 + \frac{zK''(z)}{K'(z)} = \frac{1 - 4\eta z + \eta^2 z^2}{(1-\eta z)(1+\eta z)}.$$

Putting $z = e^{i\theta}$, we have

$$\Re\left(1 + \frac{zK''(z)}{K'(z)}\right) = \frac{1 - 4\eta(1-\eta^2)\cos\theta - \eta^4(1+\cos\theta)\sin^2\theta}{((1+\eta^2)^2 - (2\eta\cos\theta)^2)}. \tag{3.7}$$

Since $((1+\eta^2)^2 - (2\eta\cos\theta)^2) > 0$ for all θ and for each fixed η. Therefore, we now only need to consider the numerator in (3.7). A computation reveals that

$$N(\theta) := 1 - 4\eta(1-\eta^2)\cos\theta - \eta^4(1+\cos\theta)\sin^2\theta$$

is increasing in $0 \leq \theta \leq \pi$ (note that $N(\theta) = N(-\theta)$) with $N(\theta) \geq 0$ when $0 < \eta \leq 2 - \sqrt{3}$, while $N(\theta)$ takes negative values when $\eta > 2 - \sqrt{3}$. Hence by the definition of convexity, result follows.

Now let us consider the function

$$\psi(z) := \frac{\gamma z}{(1+\eta z)^2} = \gamma K(z), \quad \text{where} \quad \gamma > 0,$$

and introduce a related class defined as follows:

$$\mathcal{S}_\gamma(\eta) := \left\{ f \in \mathcal{A} : \left(\frac{zf'(z)}{f(z)} - 1 \right) \prec \frac{\gamma z}{(1+\eta z)^2},\ \eta \in [0,1),\ \gamma > 0 \right\}. \tag{3.8}$$

Note that if γ and η satisfies the condition $(1-\eta)^2 \geq \gamma$, then the class $\mathcal{S}_\gamma(\eta)$ reduces to a Ma-Minda subclass of univalent starlike functions. Also letting $\eta = 1/4$ and $\gamma = 25(\sqrt{2}-1)/16$, we see that the class $\mathcal{S}^*(\sqrt{1+z}) \subset \mathcal{S}_\gamma(\eta)$.

Theorem 3.11. *Let $f \in \mathcal{S}_\gamma(\eta)$ and $\eta \in [0, 2-\sqrt{3}]$. Then*

$$-\kappa(-r) \leq |f(z)| \leq \kappa(r),$$

where

$$\kappa(z) := z \exp\left(\frac{\gamma z}{(1+\eta z)^2} \right).$$

Proof. Since $\psi(z) = \gamma K(z)$, using Lemma 3.1, we see that for $|z| = r$, $\psi(-r) \leq \Re\psi(z) \leq \psi(r)$. Also, we have $\kappa(z) = z \exp \int_0^z (\psi(t)/t)dt$. Hence, the result follows from Theorem 1.4.

Using Lemma 3.1, we also obtain that $\Re\psi(z) \geq \psi(-r)$ for all $\eta \in [0,1)$ which implies $-\kappa(-r) \leq |f(z)|$. So we have the following results:

Corollary 3.12 (Radius of starlikeness)**.** *Let $f \in \mathcal{S}_\gamma(\eta)$, $\gamma > 0$ and $\eta \in [0,1)$. Then f is starlike (univalent) of order $\alpha \in [0,1)$ inside the disk $|z| < r_0$, where r_0 is the smallest positive root of the equation*

$$(1-\alpha)\eta^2 r^2 - (2(1-\alpha)\eta + \gamma)r + (1-\alpha) = 0.$$

Corollary 3.13 (Koebe-radius)**.** *Let $f \in \mathcal{S}_\gamma(\eta)$ and $\eta \in [0,1)$. Then either f is a rotation of κ or*

$$\left\{ w \in \mathbb{C} : |w| \leq -\kappa(-1) = \exp\left(\frac{-\gamma}{(1-\eta)^2} \right) \right\} \subset f(\Delta).$$

Remark 3.4. Let $F_\kappa(z) := \kappa(z)/z = \exp(\gamma z/(1+\eta z)^2)$. We see that for $\eta = 0$ and $\gamma \leq 1$, F_κ is clearly convex. So consider $0 < \eta < 1$. After some calculations, we obtain that

$$G(z) := 1 + \frac{zF_\kappa''(z)}{F_\kappa'(z)} = \frac{\eta^4 z^4 + (2\eta^3 + \gamma\eta^2)z^3 - (6\eta^2 + 2\eta\gamma)z^2 + (\gamma - 2\eta)z + 1}{(1+\eta z)^3(1-\eta z)}.$$

Now for $z = e^{i\theta}$, the denominator of the real part of G is $(1+\eta^2 - 2\eta\cos\theta)(1+\eta^2 + 2\eta\cos\theta)^3 > 0$, since $(1-\eta)^2 > 0$ and therefore, it suffices to consider the numerator. After a rigorous computation, we find that numerator of the real part of G is non negative if and only if $0 < \gamma < 1$ and $0 < \eta \leq \eta_0$, where η_0 (depends on γ) is the smallest positive root of the equation

$$(1-\gamma) + (3\gamma - 10)\eta^2 + 12\eta^3 + (8 - 3\gamma)\eta^4 - 16\eta^5 + (2+\gamma)\eta^6 + 4\eta^7 - \eta^8 = 0. \tag{3.9}$$

Hence, F_κ convex for $0 < \gamma < 1$ and $0 < \eta \leq \eta_0$.

For our next result, we need to recall the following result of Ruscheweyh and Stankiewicz [13]:

Lemma 3.2 ([13])**.** *Let the analytic functions F and G be convex univalent in Δ. If $f \prec F$ and $g \prec G$, then*

$$f * g \prec F * G \quad (z \in \Delta).$$

Theorem 3.14. *Let $\eta \in [0, 2-\sqrt{3}]$. If f belongs to the class $\mathcal{S}_\gamma(\eta)$, then*

$$\frac{f(z)}{z} \prec F_\kappa(z), \quad (z \in \Delta)$$

where $F_\kappa(z) = \kappa(z)/z$ is the best dominant and κ as defined in Theorem 3.11.

Proof. Let $f \in \mathcal{S}_\gamma(\eta)$, then

$$\phi(z) := \frac{zf'(z)}{f(z)} - 1 \prec \psi(z). \tag{3.10}$$

It is well-known that the function

$$g(z) = \log\left(\frac{1}{1-z}\right) = \sum_{n=1}^{\infty} \frac{z^n}{n} \in \mathcal{C},$$

where $\mathcal{C}$ is the usual class of normalized convex(univalent) function and thus for $f \in \mathcal{A}$, we get

$$\phi(z) * g(z) = \int_0^z \frac{\phi(t)}{t} dt \quad \text{and} \quad \psi(z) * g(z) = \int_0^z \frac{\psi(t)}{t} dt. \tag{3.11}$$

From Lemma 3.1, we see that ψ is convex for $\eta \in [0, 2-\sqrt{3}]$. Thus applying Lemma 3.2 in (3.10), we get

$$\phi(z) * g(z) \prec \psi(z) * g(z). \tag{3.12}$$

Now from (3.11) and (3.12), we obtain

$$\int_0^z \frac{\phi(t)}{t} dt \prec \int_0^z \frac{\psi(t)}{t} dt,$$

which implies that

$$\frac{f(z)}{z} := \exp \int_0^z \frac{\phi(t)}{t} dt \prec \exp \int_0^z \frac{\psi(t)}{t} dt =: \frac{\kappa(z)}{z}.$$

This completes the proof.

Corollary 3.15. *Let $0 < \gamma < 1$ and $0 < \eta \leq \min\{2-\sqrt{3}, \eta_0\}$, where η_0 is the least positive root of the equation (3.9) and also let $0 < \gamma \leq \pi/2$ when $\eta = 0$. If $f \in \mathcal{S}_\gamma(\eta)$, then f satisfies the sharp inequality*

$$\left|\arg \frac{f(z)}{z}\right| \leq \max_{|z|=r} \arg \exp\left(\frac{\gamma z}{(1+\eta z)^2}\right).$$

Proof. Let $F_\kappa(z) := \kappa(z)/z = \exp(\gamma z/(1+\eta z)^2)$ which is symmetric about the real axis. From Theorem 3.14, have $f(z)/z \prec F_\kappa(z)$ for $0 \leq \eta \leq 2-\sqrt{3}$. Since for $\eta = 0$, $\Re F_\kappa(z) > 0$ if and only $\gamma \leq \pi/2$. The result is obvious. Now from Remark 3.4, we

see that if $0 < \gamma < 1$ and $0 < \eta \leq \min\{2-\sqrt{3}, \eta_0\}$, where η_0 is the least positive root of the equation (3.9) then F_κ is convex which implies

$$\Re F_\kappa(z) \geq \exp\left(\frac{-\gamma}{(1-\eta)^2}\right) > 0,$$

and F_κ is also a Caratheódory function in this case. Hence the result follows.

Now using Theorem 3.11, Remark 3.4 and Theorem 3.14, we obtain the following result:

Theorem 3.16. *Let $f \in \mathcal{S}_\gamma(\eta)$, then*

$$\Re\left(\frac{f(z)}{z}\right) \leq \exp\left(\frac{\gamma r}{(1+\eta r)^2}\right) \quad \textit{for} \quad \eta \in [0,1)$$

and

$$\min_{|z|=r} \exp\left(\frac{\gamma z}{(1+\eta z)^2}\right) \leq \Re\left(\frac{f(z)}{z}\right) \quad \textit{for} \quad \eta \in [0, 2-\sqrt{3}].$$

In partcular, if $0 < \gamma < 1$ and $0 < \eta \leq \min\{2-\sqrt{3}, \eta_0\}$, where η_0 is the least positive root of the equation (3.9), then

$$\exp\left(\frac{-\gamma r}{(1-\eta r)^2}\right) \leq \Re\left(\frac{f(z)}{z}\right).$$

The result is sharp.

We conclude this paper by introducing the following three new subclasses of $\mathcal{F}(\psi)$:

$$\mathcal{T} := \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} - 1 \prec \log(1-z) \right\},$$

which means $zf'(z)/f(z) \in \{w \in \mathbb{C} : |\exp(w-1) - 1| < 1\}$,

$$\mathcal{S}_p := \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} \prec 1 - \left(\log\frac{1+\sqrt{z}}{1-\sqrt{z}}\right)^2 \right\},$$

or equivalently $zf'(z)/f(z) \in \{w \in \mathbb{C} : |1-w| < \Re((1-w)+\pi^2)\}$, a parabola with opening in left half plane and

$$\mathcal{L}(\beta) := \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} - 1 \prec \frac{z}{\cos(\beta z)},\ \beta \in [0,1] \right\}.$$

The above new classes are still open to study. Also see figure 1. Note that for the classes $\mathcal{T}$ and $\mathcal{L}(\beta)$, the function f_0 defined in (2.5) takes the respective particular form

$$f_{\mathcal{T}}(z) := z\exp(-Li_2(z)),$$

where

$$-\int_0^z \frac{\log(1-t)}{t}dt = \sum_{n=1}^{\infty}\frac{z^n}{n^2} =: Li_2(z)$$

known as dilogarithm function and

$$f_{\mathcal{L}}(z) := z\exp\int_0^z \frac{1}{\cos\beta t}dt = z(\sec\beta z + \tan\beta z)^{1/\beta}), \quad \beta \neq 0.$$

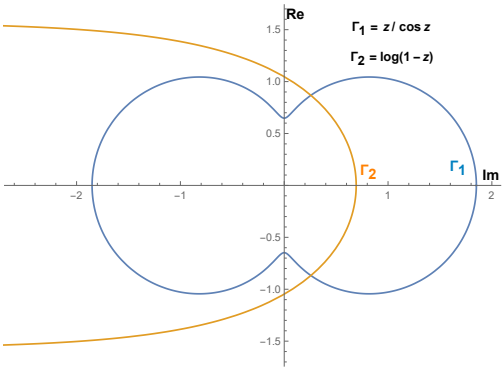


FIGURE 1. Boundary curves of the functions $z/\cos z$ and $\log(1-z)$

CONCLUSION

It is interesting to observe that even in the class $\mathcal{F}(\psi)$, functions may not be univalent. But with the conditions on the bounds for the real part of ψ, a similar result holds as obtained by Ma-Minda [8] which is quiet important to obtain the Koebe domain. From Remark 3.2 and Remark 3.4, we also note that the function $f_0(z)/z$, where f_0 as defined in (2.5) behaves quite differently in the particular classes.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

[1] ALI, R. M.—JAIN, N. K.—RAVICHANDRAN, V.: *Bohr radius for classes of analytic functions*, Results Math. **74** (2019), Art. 179, 13 pp.

[2] BHOWMIK, B.—DAS, N.: *Bohr phenomenon for subordinating families of certain univalent functions*, J. Math. Anal. Appl. **462** (2018), 1087–1098.

[3] CHO, N. E.—KUMAR, S.—KUMAR, V. —RAVICHANDRAN, V.: *Differential subordination and radius estimates for starlike functions associated with the Booth lemniscate*, Turkish J. Math. **42** (2018), 1380–1399.

[4] GOEL, P.—SIVAPRASAD KUMAR, S.: *Certain class of starlike functions associated with modified sigmoid function*, Bull. Malays. Math. Sci. Soc. **43** (2020), 957–991.

[5] KANAS, S.—WIŚNIOWSKA, A.: *Conic domains and starlike functions*, Rev. Roumaine Math. Pures Appl. **45** (2000), 647–657.

[6] KARGAR, R.—EBADIAN, A.—SOKÓŁ, J.: *On Booth lemniscate and starlike functions*, Anal. Math. Phys. **9** (2019), 143–154.

[7] KARGAR, R.—EBADIAN, A.—TROJNAR-SPELINA, L.: *Further results for starlike functions related with Booth lemniscate*, Iran. J. Sci. Technol. Trans. A Sci. **43** (2019), 1235–1238.

[8] MA, W.C.—MINDA,D.: *A unified treatment of some special classes of univalent functions*, *Proceedings of the Conference on Complex Analysis (Tianjin, 1992)*, Conf. Proc. Lecture Notes Anal., I Int. Press, Cambridge, MA., 157–169.

[9] MASIH, V. S.—EBADIAN, A.—YALÇIN, S.: *Some properties associated to a certain class of starlike functions*, Math. Slovaca **69** (2019), 1329–1340.

[10] NAJMADI, P.—NAJAFZADEH, SH.—EBADIAN, A.: *Some properties of analytic functions related with Booth lemniscate*, Acta Univ. Sapientiae Math. **10** (2018), 112–124.

[11] PIEJKO, K.—SOKÓŁ, J.: *Hadamard product of analytic functions and some special regions and curves*, J. Inequal. Appl. **2013**, (2013), Art. 420, 13 pp.

[12] PIEJKO, K.—SOKÓŁ, J.: *On Booth lemniscate and hadamard product of analytic functions*, Math. Slovaca. **65** (2015), 1337–1344.

[13] RUSCHEWEYH, S.—STANKIEWICZ, J.: *Subordination under convex univalent functions*, Bull. Polish Acad. Sci. Math. **33** (1985), 499–502.

[14] SHARMA, P.—RAINA, R. K.—SOKÓŁ, J.: *Certain Ma-Minda type classes of analytic functions associated with the crescent-shaped region*, Anal. Math. Phys. **9** (2019), 1887–1903.

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India

E-mail address: spkumar@dce.ac.in

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India

E-mail address: gangania.m1991@gmail.com

Performance Analysis of a Novel Hetero-material InAs/GaAs Junctionless TFET

Samriti Sharma
*Delhi Technological Univer*sity
New Delhi, India
Sharmasamriti72@gmail.com

Rishu Chaujar
*Delhi Technological Univer*sity
New Delhi, India
chaujar.rishu@dtu.ac.in

***Abstract*—In this work, we proposed and investigated the performance features of a hetero-material junctionless tunnel field-effect transistor (HM-JLTFET) by incorporating□□-compound semiconducting materials, InAs having low work function, in the source section and GaAs having higher work function, in the channel and drain sections, respectively. The transfer characteristics, subthreshold slope, leakage current, and current switching ratio of the proposed device are also compared with the conventional TFET device having the same dimensions as that of HM-JLTFET. The simulated results reflect the capability of the proposed device to suppress the performance degradation factors like leakage current and power consumption along with improving the current driving capability and fast switching at low gate voltage.**

Keywords—Hetero-material, Junctionless, TFET, Band-gap, Indium arsenide, Polar gate.

I. INTRODUCTION

Due to the thermionic emission as the fundamental mechanism of MOSFET, the performance of MOSFET degrades comprising of subthreshold slope value larger than 60 mV/decade and high leakage current. Many other conventional devices like CMOS (the combination of nMOS and pMOS), High Electron Mobility Transistor (HEMT), Negative capacitance FET, and Heterojunction Bipolar Transistor (HBT) appear to be promising in one form of challenges while degrades in other technological challenges. The nonstop scaling down of CMOS technology in the past few years has led to many operational enhancements in terms of high speed and packing density; however, it becomes very challenging to fabricate ultra-shallow junctions at such small dimensions [1]-[3]. Various techniques like ion implantation and annealing are required in order to fabricate ultra-shallow junctions with heavy doping, which, however, suffers from random dopant fluctuations [4]-[5].

Along with such issues, the downscaling of MOSFET lead to static power dissipations and high leakage current [6]-[7]. TFETs are the devices, which overcome the issues of higher leakage current and fast switching along with lower subthreshold slope. The underlying mechanism of TFET device is the band to band quantum excavation of electrons through the barrier, which results into high ON current and high I_{ON}/I_{OFF} current switching ratio [8]. In the past few decades, many research papers have highlighted the importance of TFET; however, low ON-current and fabrication of metallurgical junctions has been a complicated task. In order to meet the requirements of modern technology, junctionless TFETs have come into the picture [3]. Junctionless TFETs integrate the advantages of junctionless FETs and TFETs. Junctionless FET shows high ON current and has less variability as compared to MOSFETs as it does not require any metallurgical junction. TFETs show steeper SS, however, suffer from low ON-current [3], [9]. JLTFET works on the principle of inducing source and drain regions through charge plasma [3]. This paper includes the amalgamation of the concept of JLFET and TFET along with ˜ -˜ compound semiconducting materials, where, InAs is used in the source section and GaAs in the channel and drain sections, respectively.

II. DEVICE ARCHITECTURE AND SIMULATION SETUP

Fig. 1(a) illustrates the pictorial outlook of our proposed device double gate hetero-materials junctionless TFET (HM-JLTFET). The configuration of the device is a lateral n-type junctionless TFET consisting of a uniformly doped substrate with concentration of 1×10^{19} cm^{-3}. Our device is based on charge plasma concept using two isolated gates as the Polar gate and Control gate, where the Control gate has lower work function than Polar gate. The primary purpose of these isolated gates is to convert a highly doped $n^+ - n^+ - n^+$ substrate into the conventional $p^+ - i - n^+$structure. The device structure parameters are mentioned in Table I. The work functions of the control gate and polar gate are 4.5 eV and 5.93 eV, respectively.

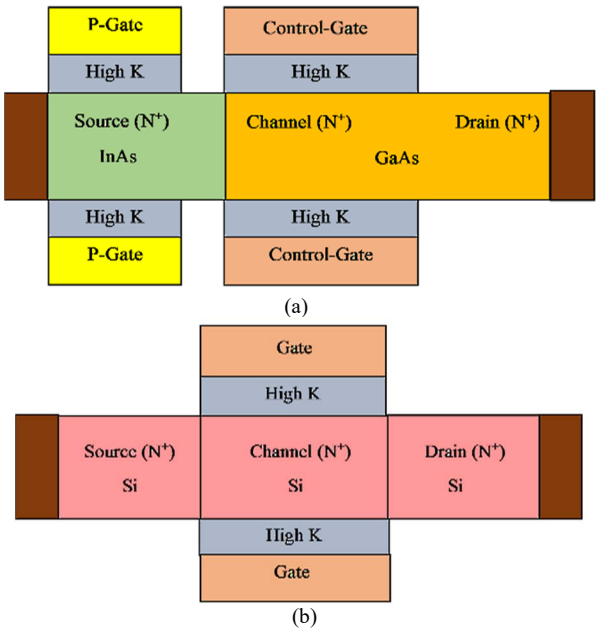


Fig. 1. Pictorial outlook of (a) Hetero-material junctionless TFET (HM-JLTFET) and (b) Conventional TFET

978-1-7281-1933-5/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:28:16 UTC from IEEE Xplore. Restrictions apply.

TABLE I. DEVICE STRUCTURE PARAMETERS

HM-JLTFET		
Device Parameter	***Value***	***units***
Polar-Gate (PG) length	15	nm
Control-Gate (CG) length	20	nm
Oxide Thickness	2	nm
Work function of PG	5.93	eV
Work function of CG	4.5	eV
Isolation between PG and CG	2	nm

The channel, source, and drain has length of 20 nm, 15 nm, and 15 nm, respectively. A polar gate is applied over the source region to induce p-type source, having length of 15 nm. A high-k gate dielectric, Al_2O_3 is used, having width of 2 nm. The body thickness is 3 nm. The results of the proposed HM-JLTFET are also compared with the conventional TFET keeping the dimensions same for both devices. For conventional device, Si is used as the substrate and gate material has the work function of 4.7 eV. The source region is doped with concentration of 5×10^{19} cm^{-3}, channel region is 1×10^{15} cm^{-3} and drain region is 1×10^{19} cm^{-3}.

III. SIMULATION METHODOLOGY

Simulations are performed in the SILVACO-ATLAS 2-D simulator [11]. Since TFET works on the mechanism of band to band tunneling through the tunneling barrier, so non-local Band-to-Band tunneling model is applied. The non-local models are independent of the electric field at every individual mesh node of the simulated structure. It relies on the band structure of the device and also the spatial variation of the bands. Band gap narrowing (BGN) and Auger recombination model are also applied. To include the effects of traps, trap assisted tunnelling model is also involved. Owing to the higher impurity concentration in the channel region, SRH recombination model is used. Fermi Dirac statistics and CVT model (Lombardi) are also invoked. The CVT model includes the parallel, perpendicular electric field mobility, temperature-dependent mobility, and concentration-dependent mobility. As the body thickness is 3 nm, so quantum confinement model given by Hansch [12], is also enabled.

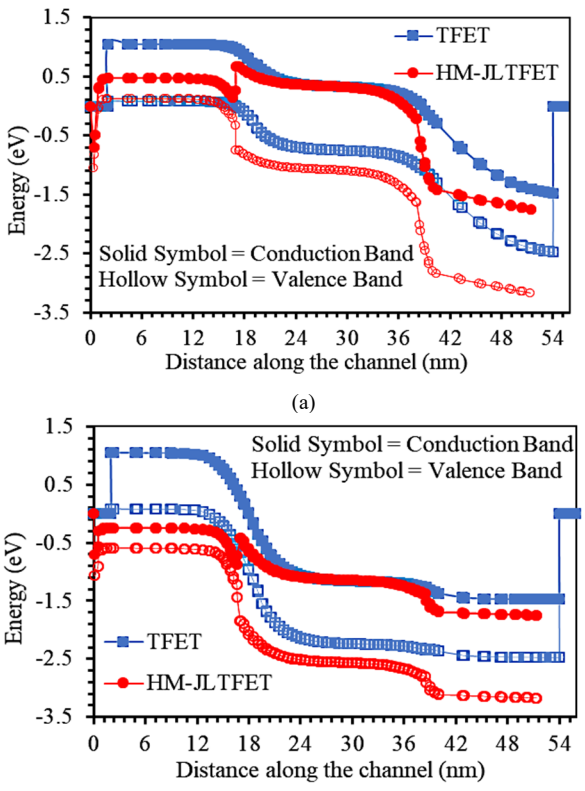


Fig. 2. Energy diagram (a) OFF state and (b) ON state of HM-JLTFET and TFET.

IV. RESULTS AND DISCUSSION

Fig. 2 (a) presents the band diagram of hetero-material junctionless tunnel FET and conventional TFET for OFF state ($V_{GS} = 0.0\,V, V_{DS} = 1.5\,V$). As explained before, HM-JLTFET consists of InAs as the source and GaAs as the channel material. The energy band gap of InAs is 0.35 eV and GaAs has 1.43 eV energy gap. In the OFF state, the excavating barrier thickness of HM-JLTFET is much wider than the conventional TFET due to the incorporation of GaAs as the channel and drain material, which has higher energy band gap than Si. Owing to the wider excavating barrier

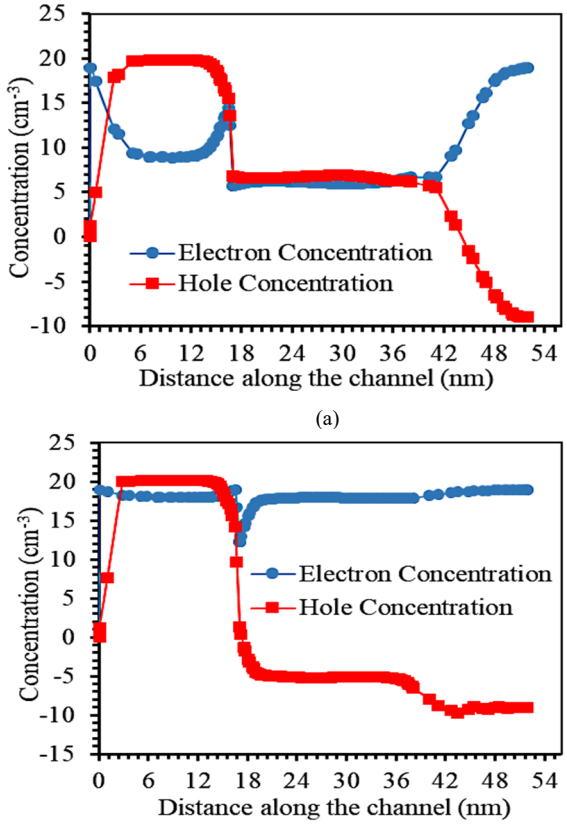


Fig. 3. Carrier concentration of HM-JLTFET for (a) OFF state and (b) ON state.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:28:16 UTC from IEEE Xplore. Restrictions apply.

thickness at the source/channel (S/C) intersection, only a few electrons can tunnel through the barrier to reach across the drain region in HM-JLTFET. Therefore, only a low leakage current flow through the tunneling barrier of the device. In TFET, Since the tunneling barrier thickness is smaller in comparison to HM-JLTFET, so there is a massive flow of electrons through the S/C boundary giving rise to high leakage current flow through the device. When gate biasing is applied ($V_{GS} = 1.5\,V, V_{DS} = 1.5\,V$), as exposed in Fig. 2(b), the conduction band of HM-JLTFET bends downward while the valence band is fetched up as a consequence of electric field produced by the control gate at the S/C boundary. A large number of electrons excavate from the source to the channel causing high ON current flow in the device. The Tunneling barrier width is more significant in HM-JLTFET due to which, the number of electrons tunneling through the barrier is smaller as compared to TFET. Hence the ON current of HM-JLTFET is 24.3 µA/µm whereas the ON current of TFET is 29.7 µA/µm.

Fig. 3(a) and 3(b) represents the carrier concentration of HM-JLTFET in ON and OFF state. In the ON state, the polar gate with work function 5.93 eV induces the p-type source region and increases the hole concentration from 1×10^{19} cm^{-3} to 1×10^{20} cm^{-3}. The control gate having the work function 4.5 eV induces the intrinsic channel region by rising the electron concentration from 1×10^{7} cm^{-3} to 1×10^{17} cm^{-3} on applying the gate voltage of 1.5 V.

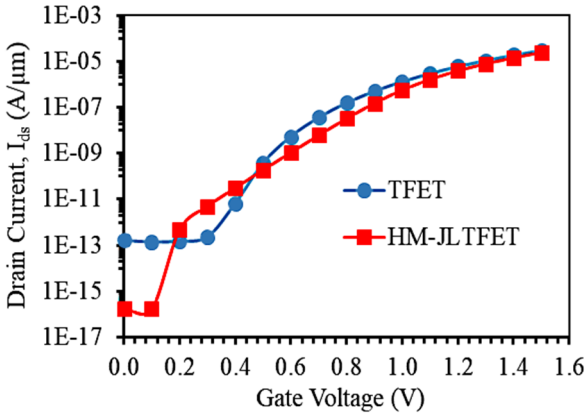


Fig. 4. Transfer characteristics of HM-JLTFET and TFET at V_{GS}=1.5V, V_{DS}=1.5V

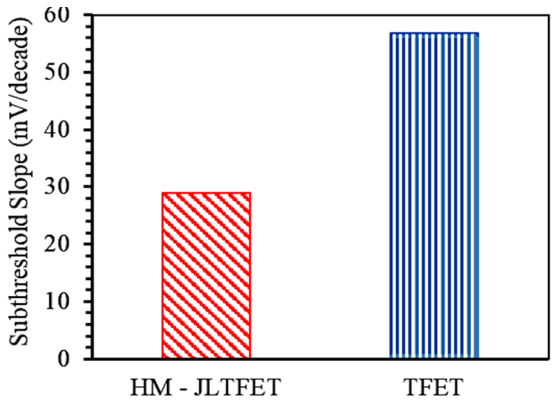


Fig. 5. Comparison plot of subthreshold slope (SS) of HM-JLTFET and TFET.

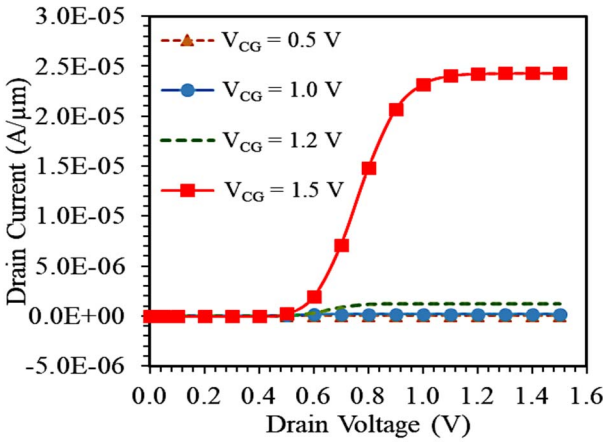


Fig. 6. Output characteristics of HM-JLTFET for different control gate voltages.

The drain region remains n-type by electron concentration of 1×10^{19} cm^{-3}. The charge plasma concept used in HM-JLTFET converts a highly doped $n^+ - n^+ - n^+$ substrate into the conventional $p^+ - i - n^+$ structure. On applying the biasing to the gate and drain electrodes , the source region attains the p-type charge concentration of 1×10^{20} cm^{-3} and drain region attains the n-type charge concentration of 1×10^{19} cm^{-3} whereas the channel remains intrinsic with concentration of 1×10^{17} cm^{-3}.

Fig. 4 represents the comparison plot of transfer characteristics of HM-JLTFET and conventional TFET. From Fig. 4, it is inferred that the OFF current in HM-JLTFET is about 1.7×10^{-16} A/µm and the ON current comes out to be 2.43×10^{-5} A/µm. The OFF current of conventional TFET rises to 1.28×10^{-12} A/µm and the ON current reaches to 2.96×10^{-5} A/µm. In HM-JLTFET, on the application of III-IV compound semiconducting hetero-materials in the source and channel, the tunnel barrier thickness increases in the OFF state and hence the movement of electrons through the tunnel barrier is prohibited, leading to lower OFF state current.

In the case of conventional TFET, Si is used as the channel material. Having smaller energy band gap than GaAs, it increases the movement of electrons through the tunnel barrier, and hence the leakage current increases to a higher rate. Fig. 5 represents the comparison plot of the subthreshold slope (SS) of HM-TFET and TFET. HM-TFET shows the steeper subthreshold slope of 29 mV/ decade, whereas the conventional TFET shows larger subthreshold slope of 57 mV/decade. Fig. 6 represents the output characteristics of our proposed device for various control gate voltages. It depicts the exponential and saturation section of the device process.

For the exponential section, the current relies linearly on the drain to source voltage, whereas for the saturation section, the excavating width depends more on the gate bias and shows no dependence on the drain bias. As the control gate voltage increases from 0.5 V to 1.5 V, the tunneling barrier width among the valence band at the source and conduction band at the drain decreases. Hence electrons can easily flow from the source to channel region. It improves the drain current of HM-JLTFET.

Fig. 7 represents the comparative plot of current switching ratio of HM-JLTFET and TFET. All the result features of the HM-JLTFET and conventional TFET are displayed in Table II.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:28:16 UTC from IEEE Xplore. Restrictions apply.

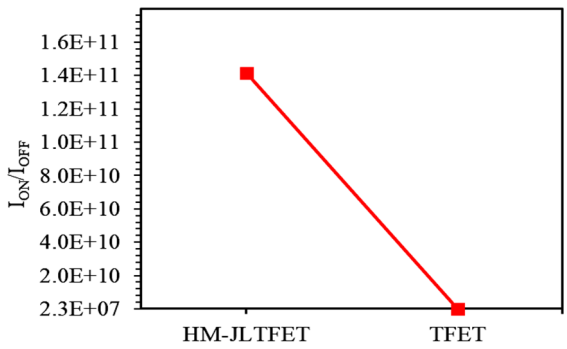


Fig. 7. Comparison plot of the I_{ON}/I_{OFF} of HM-JLTFET and TFET.

The higher I_{ON}/I_{OFF} of HM-JLTFET (1.4×10^{11}) makes it suitable for fast switching applications as compared to TFET having I_{ON}/I_{OFF} ratio of 2.3×10^{7}.

TABLE II RESULT CHARACTERISTICS

Parameter	***HM-JLTFET***	***TFET***
I_{OFF} (A/μm)	1.72×10^{-16}	1.28×10^{-12}
I_{ON} (μA/μm)	24	29
I_{ON}/ I_{OFF}	1.4×10^{11}	2.3×10^{7}
SS (mV/decade)	29	57

V. CONCLUSION

In this paper, we proposed a double gate hetero-material junctionless TFET based on compound semiconducting material InAs and GaAs as the source and channel material, respectively, by invoking the band gap engineering and junctionless engineering. We also compared the device performance with conventional TFET. The proposed device provides steeper subthreshold slope, which reduces to almost half (49%) as that of conventional TFET on the application of compound semiconducting materials in the source and channel region. HM-JLTFET shows very low leakage current, which is reduced to 10^4 times as that of TFET. HM-JLTFET shows fast switching I_{ON}/I_{OFF} ratio, which is 10^4 times higher as compared to conventional TFET and makes it appropriate for high speed operations and low leakage circuits along with more facile fabrication due to the hetero-junctionless technology.

ACKNOWLEDGMENT

The authors would like to acknowledge the Microelectronics Research Lab, Department of Applied Physics, Delhi Technological University, to carry out this work.

REFERENCES

[1] Goswami, Yogesh, Bahniman Ghosh, and Pranav Kumar Asthana. "Analog performance of Si junctionless tunnel field effect transistor and its improvisation using III–V semiconductor." *RSC Advances* 4.21 (2014): 10761-10765.

[2] Asthana, Pranav Kumar, et al. "High-speed and low-power ultradeep-submicrometer III-V heterojunctionless tunnel field-effect transistor." *IEEE Transactions on Electron Devices* 61.2 (2014): 479-486.

[3] Ghos h, Bahniman, and Mohammad Waseem Akram. "Junctionless tunnel field effect transistor." *IEEE electron device letters* 34.5 (2013): 584-586.

[4] Niga m, Kaushal, et al. "A new approach for design and investigation of junction-less tunnel FET using electrically doped mech anism." *Superlattices and Microstructures* 98 (2016): 1-7.

[5] Dam rongplasit, Nattapol, Sung Hwan Kim, and Tsu-Jae King Liu. "Study of random dopant fluctuation induced variability in the raised-Ge-source TFET." *IEEE Electron Device Letters* 34.2 (2013): 184-186.

[6] Tirkey, Sukeshni, Dharmendra Singh Yadav, and Dheeraj Sharma. "Controlling ambipolar current of dopingless tunnel field-effect transistor." *Applied Physics A* 124.12 (2018): 809.

[7] Vadizadeh, Mahdi. "Performance estimation of junctionless field effect diode." *Applied Physics A* 125.8 (2019): 495.

[8] Kim, Sang Wan, et al. "Demonstration of L-shaped tunnel field-effect transistors." *IEEE transactions on electron devices* 63.4 (2015): 1774-1778.

[9] Raad, Bhagwan Ram, et al. "Drain work function engineered doping-less charge plasma TFET for ambipolar suppression and RF performance improvement: A proposal, design, and investigation." *IEEE Transactions on Electron Devices* 63.10 (2016): 3950-3957.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:28:16 UTC from IEEE Xplore. Restrictions apply.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/344027034

PERFORMANCE MEASUREMENT OF TQM USING INTEGRATED FUZZY-AHP

Article · September 2020

DOI: 10.24247/ijmperdjun20201008

CITATIONS
0

READS
12

2 authors, including:



Rakesh Kumar
Delhi Technological University
8 PUBLICATIONS **1** CITATION

SEE PROFILE

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

Population Statistics View project

Performance Measurement View project

All content following this page was uploaded by Rakesh Kumar on 01 September 2020.

The user has requested enhancement of the downloaded file.

International Journal of Mechanical and Production Engineering Research and Development (IJMPERD)
ISSN(P): 2249–6890; ISSN(E): 2249–8001
Vol. 10, Issue 3, Jun 2020, 10543–10562
© TJPRC Pvt. Ltd.

TRANS STELLAR
•Journal Publications •Research Consultancy

PERFORMANCE MEASUREMENT OF TQM USING INTEGRATED FUZZY-AHP

RAKESH KUMAR & R.S. MISHRA

Delhi Technological University, Delhi-110042

ABSTRACT

Performance measurement is regarded as one of the most important and critical tasks in the course of any organization. Measuring performance is the assessment of the possible impact that a proposed plan or project may have on the organizational system, together consisting of the financial, non-financial and operational aspects. The aim of this paper is to propose a performance measurement methodology based on an integrated Fuzzy-AHP approach in the context of total quality management (TQM). In the proposed methodology the criteria weights are generated by a fuzzy-AHP procedure. The fuzzy set theory is a perfect means for modeling uncertainty or imprecision arising from human mental phenomena. The usage of fuzzy sets in describing uncertainties and vagueness in different critical success factors of TQM simplifies the complex structure of performance assessment.

KEYWORDS: Performance Measurement, TQM, Multicriteria, Fuzzy, AHP

Received: Jun 08, 2020; **Accepted:** Jun 28, 2020; **Published:** Sep 01, 2020; **Paper Id.:** IJMPERDJUN20201008

Original Article

INTRODUCTION

Performance measurement often has to deal with latent variables which were difficult to identify and define also components that may involve both quantitative and qualitative factors. Performance measurement is a complex task not only because of its wide scope but also because of the wide range of attributes that bear on its assessment, Kumar, R. et al. (2019).In terms of scope, an assessment may cover subjective and vague decisions that may affect various interest groups or stakeholders each with their own demands and needs. In view of these difficulties, methods based on fuzzy logic may be quite useful in undertaking difficult assessment procedures. The fuzzy set theory was introduced by Zadeh (1965) to express the linguistic terms in decision-making process in order to resolve the vagueness, ambiguity and subjectivity of human judgment. A fuzzy method stands very helpful to tackle precisely with complex and ill-defined performance measurement decision making problems. Very few organizations categorized themselves as having fairly TQM-driven performance management system, they were quite positive – if not yet revealing conclusive evidence – of understanding the influence of system factors on individual performance in a quality-driven context, Soltani,E. et al. (2005).G. G. Dess et al. (1984) addressed in his research that research which incorporates organizational performance (OP) must address two basic issues: (1) selection of a conceptual framework from which to define OP and (2) identification of accurate, available measures that operationalize OP. In addition, collecting data and information to coordinate and support the procedure of taking action and making decisions throughout the organization is the reason for performance measurement (Van Schalkwyk, 1998).

The subsystems like Just-in-Time (JIT), Total Productivity System (TPS), Total Productive Maintenance (TPM), Supply Chain Management (SCM), Lean Manufacturing (LM) System and many more were incorporated to the organizational system as per their fit to achieve the benchmarked performance or the targeted level of

performance. Vanichchinchai, A., and Igel, B. (2011) mentioned that to achieve JIT delivery, quality of the whole internal operations as well as that of the external partners must be assured. Scannell et al. (2000) found that JIT purchasing significantly associated with flexibility, conformance to specification and low production cost performance. Total Quality Management (TQM) is one of them most comprehensive management approach to ensure the higher levels of performance of the organization. TQM is not limited only to ensure the quality of products but it is equally important to vigil the performance of organization and also covers wider aspects of quality product and services. British Standard Institution defines; TQM consists of a "management philosophy and company practices which aim to harness the human and material resources of a corporation within the best way to achieve the objectives of the organization". Definition of TQM by Feigenbaum's (1988) encompasses every aspect of the TQM cycle including, customer focus, process planning, process management, and process improvement. Feigenbaum's (1988) states that "a total quality system is defined as one which embraces the whole cycle of customer satisfaction from the interpretation of his requirements prior to the ordering stage, through to the supply of a product or service at an economical price and on to his perception of the product after he has used it over an appropriate period of time".

Instead of investigating the gap based on a specific theoretical view, this study is an exploratory and quantitative effort focused on the identification of sources of the gap. As objective of this study, an integrated fuzzy AHP methodology is proposed for Performance measurement of TQM system. In the proposed methodology, the weights of the assessment criteria are determined by a fuzzy AHP procedure. The rest of the paper is organized as follows: In Section 2, a literature review about multiple attribute performance measurement is briefly given. In the third section, an integrated fuzzy AHP methodology is presented. In Section 4, following the determination of the selection criteria and alternatives, the proposed methodology is applied to a industrial problem. Finally, in the last section, conclusions and suggestions for further study are given. In sum, this study provides useful insights to support the decision making for performance measurement of companies to make strategic and resource allocations for developing high quality products to improve operational performance.

LITERATURE REVIEW

Performance measurement is a critical factor for the effective management. This may stem back from the fact that without measuring something, it is difficult to improve it, Ismail Salaheldin, S. (2009) strongly recommended. Therefore, improving the organizational performance requires identifying and measuring the impact of TQM practices on it (Demirbag et al., 2006; Koh et al., 2007).Performance Measurement (PM) is the process of identifying, predicting, evaluating and mitigating the critical success factors and other relevant effects of development proposals prior to major decisions being taken. Measuring performance is an important practice for ensuring that the likely effects of new development on the organizational system are fully understood and taken into account before the development is allowed to proceed. Performance appraisal, on the other hand, is defined as, 'a formal documented system for the periodic review of an individual's performance' (Moon, 1993, p. 8); or, 'the process by which data are collected and reviewed about an individual employee's past and current work behaviour and performance' (Bratton and Gold, 1999).Performance Measurement enables organizational factors to be given due weight, along with financial, non-financial or operational factors, when planning applications are being considered. If properly carried out, it benefits all those involved in the planning process. For the production planning and other strategist with quality control responsibilities, performance

measurement provides a basis for better decision making. The fundamental objectives of any organization are to establish and maintain competitive advantage and ensure strong organizational performance (Eisenhardt et al., 1998).

Performance Measurement and TQM

TQM practices have been successful in increasing organisational performance is that TQM serves as a link between an organisation's business or corporate strategy and its performance. TQM creates a system of checks and balances to ensure that an organisation's business strategy is successful (Prajogo and Sohal, 2006), the same pair author in year 2004 provided a multidimensional view of TQM practices in determining organizations' quality and innovation performance. TQM use to improve organizational efficiency and their effectiveness as well as to create tools to better manage the quality of products and services for their long-term customers. Pfeffer and Sutton (2006), argue that TQM practices emphasise evidence-based management that concentrates on creating visible and measurable improvements to processes, rather than on vague estimations of customer or client approval. A study presented by Ahire et al. (1996) about an analysis on the comparison for the Quality management in TQM firm versus non-TQM firms. Firms are classified as: high performance TQM firms, low performance TQM firms, high performance non-TQM firms and low performance non-TQM firms. Major reasons behind the failure of TQM effort viz. lack of management commitment, unrealistic expectations and time-frame and cost of TQM implementation, under-reliance on statistical methods, and failure to develop and sustain a quality-oriented culture are reported by authors. Measuring critical success factors (CSF) of TQM is an essential precondition to control the implementation process and to increase the chances of success, stated Hietschold, N. et al. (2014).Because there is no consensus on certain factors or a holistic framework, so Salaheldin (2009) stated that finding and selecting adequate measurement instruments is a major challenge to companies. The CSFs of TQM can be used to gain a better understanding of the practices of quality management, to determinethe existing quality position of an organization, to assign responsibilities within an organization and to supervise quality and improvement programs (Motwani et al.,1994).

Multiple Attribute Performance Measurement

Multi-criteria decision-making (MCDM) methods enable the decision-makers handle uncertainty, complexity and conflicting objectives. MCDM problems usually require both tangible and intangible factors to be evaluated. The main problem for intangible criteria is that the values for such criteria are often imprecisely defined for the decision-makers, and the desired value for the criteria are usually defined in linguistic terms. It is difficult to express the character and significance of criteria exactly or clearly in crisp methods Chou T.Y et al. (2008).Decision making in performance measurement problems can be complex and multifaceted due to the inherent barriers, volatility, uncertainty, complexity and ambiguity factors also. Moreover, decisions in performance measurement may involve many different stakeholders with different priorities or objectives. Many theories and methods of performance for conducting an evaluation with different approach have been applied in various organizations. These approaches include Balanced Scorecard, total production analysis, regression analysis, ratio analysis, Delphi analysis, Analytic Hierarchical Process (AHP), Data Envelopment Analysis (DEA) and others. Each method has its own basic concept, aim, advantages and disadvantages (Dessler, 2000). Which one is chosen by management or decision makers for assessing performance depends on the status and type of the organization. However, all the successful enterprises have some common features, including a specific vision, positive actions, and an effective performance evaluation. Since Bellman and Zadeh (1970) developed the theory of decision behaviour in a fuzzy environment, various relevant models were developed, and have been applied to different fields such as control engineering, artificial intelligence, management science, and Multiple Criteria Decision Making

(MCDM) among others. The concept of combining the fuzzy theory and MCDM is referred to as fuzzy MCDM (FMCDM). Liou and Tzeng (2007) considered that the attributes (criteria) might exist as dependent characteristics when they used a fuzzy integral approach to integrate the performance ratings of dependent attributes. Yildirim et al. (2019) used fuzzy analytical hierarchy method to determine the weighting scores of indicators of Logistics Performance Index (LPI). It is strongly considered that LPI is accepted as an important indicator of logistical performance. The proposed model is to evaluate the LPI of the OECD (Organisation for Economic Cooperation and Development) countries within a specific time frame and compared with the existing LPI rankings. Through a managerial view Kayakutlu and Buyukozkan (2011) present an analytical framework to measure the performance factors for 3PL companies. The factors integrating the strategic and operational targets are evaluated within a framework based on four levels: performance targets, planning activities, logistics operations, and performance attributes of logistics operations. The ANP method is used to determine the most effective performance attributes. To evaluate the performance of third-party logistics (3PLs), P. Kumar et al. (2012)with the implication of an integrated model of fuzzy analytic hierarchy process (FAHP) and technique for order preference by similarity to ideal solution (TOPSIS) used. FAHP is used to assign the weight to criteria for 3PL rating in this paper. Presented work by Jiansheng Zhangab and Wei Tana (2012) shows sets up of scientific, reasonable performance evaluation system using simply AHP which is especially suitable for the small and medium 3PLs enterprise from the five levels of financial, customers, transportation, storage and technology. In service quality, Wenhua Li et al. (2017) proposed a hybrid approach based on fuzzy AHP and 2-tuple fuzzy linguistic method to evaluate in-flight service quality. Service Quality Model (SERVQUAL)as an instrument is used and a hierarchy of the evaluation index system for in-flight service quality was constructed for the study. The ratings of sub-criteria are assessed in linguistic values which are further transformed into 2-tuples and the 2-tuple linguistic arithmetic mean operator is used to obtain the average ratings of 100 respondents. Tüysüz, F., and Şimşek, B. (2017) presented Hesitant fuzzy sets-based analytic hierarchy process method for analyzing the factors affecting the performance of the branches of a cargo company and also for prioritizing them according to their level of importance.

Satty's Nine-point scale (referred from Satty (1994))

Intensity of Importance	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective
3	Weak importance of one over other	Experience and judgment slightly favor one activity over another
5	Essential/Strong Important	Experience and judgment slightly favor one activity over another
7	Demonstrated importance	An activity is favored very strongly over another; its dominance demonstrated in practice
9	Absolute importance	The evidence favoring one activity over another is of the highest possible order of affirmation
2,4,6,8	Intermediate values between the two adjacent judgments	When compromise is needed
Reciprocal of the above non-zero	If activity i has one of the above nonzero numbers assigned to it when compared with activity j then j has the reciprocal value when compared with i	A reasonable assumption
Rationals	Ratios arising from scale	If consistency were to be forced by obtaining n numerical values to span the matrix

Figure 1

AN INTEGRATED FUZZY-AHP METHODOLOGY

Fuzzy-AHP (FAHP)

Amalgamation of fuzzy set theory with AHP generate fuzzy analytical hierarchical process (FAHP) and was defined by Huang and Wu (2005) in three steps. "The first involves using triangular fuzzy numbers to transform the ideas of experts into a positive reciprocal matrix. Second, a geometric mean method that weighs the fuzzy values for each option, with the

established hierarchical connection and finally, a membership function for each option that developed the ranking of priorities."In practice, AHP has two basic applications (Cheng and Li, 2002). First, the traditional use of it is to assign weights to a set of pre-determined elements (e.g criteria, factors) and second, make decision(s) out of several alternatives. When a hierarchy has only three levels, the AHP method is equivalent to the classical Simple Additive Weighting (SAW) method (which is also known as weighted linear combination or scoring methods is a simple and most often used multi attribute decision technique). We feel that the AHP method is worth utilizing when there are more than three levels in a hierarchy. For the determination of the relative importance of evaluation criteria, fuzzy-AHP can be used since it is based on pair-wise comparisons and allows the utilization of linguistic variables. From paired comparisons a relative scale of measurement is derived. Although the pair-wise comparison approach is demanding in terms of solicited input from the experts, it offers maximum insight, particularly in terms of assessing consistency of the experts' judgment. This technique is ideal for closer examination of a selected set of performance measurement criteria in total quality management context. Saaty (1980, 1994) listed several benefits of AHP. First, it helps to decompose an unstructured problem into rational decision hierarchy (similar to decision tree). Second, it can elicit more information from the experts or decision makers by employing the pair-wise comparison of individual groups of elements. Third, it sets the computations to assign weights to the elements. Fourth, it uses the consistency measures to validate the consistency of the rating from the experts and decision makers. It is, therefore, argued to be composed of both qualitative and quantitative substances. Saaty, T. L. (1990) stated that, "The most effective way to rationalize judgments is to take a pair of elements and compare them on a single property without concern for other properties or other elements". Pair-wise comparisons in the AHP assume that the decision-maker can compare any two elements $E_i E_j$, at the same level of the hierarchy and provide a numerical value ij a for the ratio of their importance. If the element E_i is preferred to E_j, then; $a_{ij} > 1$. Correspondingly, the reciprocal property; $a_{ji} = 1/a_{ij}$; $a_{ij} > 0$, for $i = 1,2,...,n$ and $j=1,2,...,n$ always holds. Each set of comparisons for a level with n elements requires $n(n-1)/2$ judgments, which are further used to construct a positive reciprocal matrix of pair wise comparisons

The positive reciprocal comparison matrix of criteria weights is given as:

$$A=\begin{bmatrix} a_{11} & a_{12} & \ldots\ldots\ldots & a_{1n} \\ a_{21} & a_{22} & \ldots\ldots\ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots\ldots\ldots & a_{nn} \end{bmatrix} \quad (1)$$

The geometric mean of each row is calculated as:

$$Zi = \left[\prod_{j=1}^{n} a_{ij}\right]^{1/2} for\ i;\ j = 1, 2, \ldots .. n \quad (2)$$

The weight W_i is calculated as, $W_i = Zi / (Z_1 + Z_2 + \ldots\ldots + Z_n),\ \forall_i$--------(3)

The values in the scale of relative importance is crisp numeric values

Equal Importance	Moderate Importance	Strong Importance	Very Strong Importance	Extreme Importance	Intermediate Values	Values for Inverse Comparison
1	**3**	**5**	**7**	**9**	**2,4,6, 8**	**1/3,1/5,1/7 & 1/9**

Scaling for Fuzzy –AHP

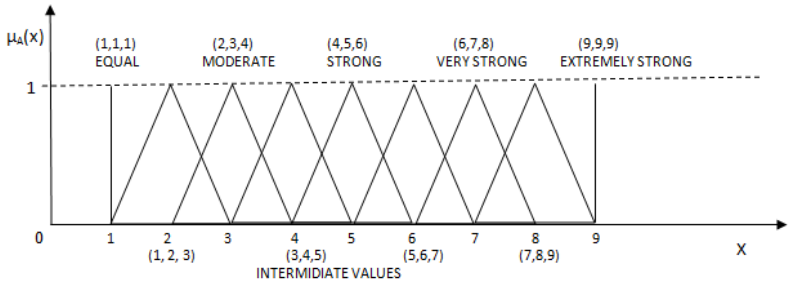


Figure 2 Fuzzy Evaluation Scale for TQM CSFs

Categories	Equally Importance	Intermediate Preference	Moderately More Important	Intermediate Preference	Strongly More Important	Intermediate Values	Very Strong More Important	Intermediate Preference	Extremely More Important
Triangular Fuzzy AHP	**1,1,1**	**1,2,3**	**2,3,4**	**3,4,5**	**4,5,6**	**5,6,7**	**6,7,8**	**7,8,9**	**9,9,9**
	1,1,1	**1/3,1/2, 1**	**1/4, 1/3,1/ 2**	**1/5,1/4,1/ 3**	**1/6,1/5,1/ 4**	**1/7,1/6,1/ 5**	**1/8,1/7,1/ 8**	**1/9,1/8,1/ 7**	**1/9,1/9,1/ 9**

To facilitate the calculation of fuzzy weights, the following arithmetic operations of triangular fuzzy numbers are presented. A triangular fuzzy number (TrFN) can be defined as A = (ℓ, m, u) where$0 \leq \ell \leq m \leq u$ as shown in Fig. 1, and its membership function is as follows:

$$\mu_A(x) = \begin{cases} \dfrac{x-\ell}{m-\ell} \text{ for } \ell \leq x \leq m; \\ \dfrac{u-x}{u-m} \text{ for } m \leq x \leq u; \\ 0, \text{ for Otherwise:} \end{cases} \quad (4)$$

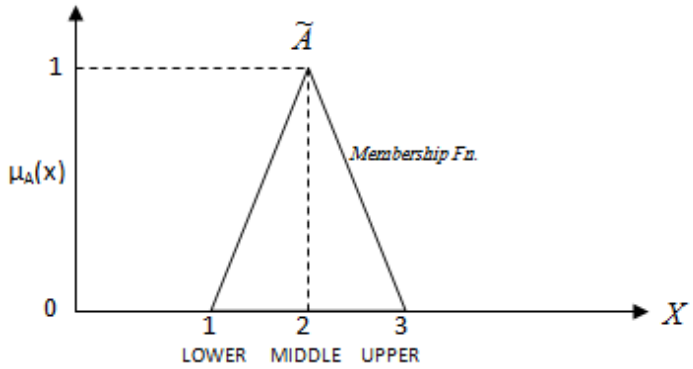


Figure 1 Membership function of TrFN

Let $u = (\ell_1, m_1, u_1)$ and $v = (\ell_2, m_2, u_2)$ be two TrFNs. Where ℓ, m, u for Lower, Middle and Upper significant membership function. The interval (ℓ, u) expresses the closure of the support of $\tilde{A}$, and the middle significant value m represents the core of $\tilde{A}$. A real number $A \in R$ can be viewed as a triangular fuzzy number.

Next to defuzzify the TrFN $m = (\ell_1, m_1, u_1)$, the following equation canbe used:

$$W_i = \frac{\ell_1 + m_1 + u_1}{3} \text{---------- (5)}$$

STEPS OF THE PROPOSED FUZZY-AHP ALGORITHM

The proposed method consists of some simple steps which have been outlined in the following sections. This method may be used for both group as-well-as individual decision-making.

Step I: Determining the Attributes. The attributes i.e. TQM drivers, enablers and capabilities are selected on the recommendation of analyst opinion. The attributes which have some sub-attributes are not included in the decision-making process. In other words, the final level of the hierarchical decision-making diagram (i.e., attributes) is involved in decision making.

Step II: Evaluate the relative importance of the criteria using pair-wise comparisons. The experts are required to provide their judgments on the basis of their knowledge and expertise. The experts' linguistic preferences are converted into triangular fuzzy numbers using Table 2.

Step III: Aggregate experts' individual preferences into group preference by applying the fuzzy triangular averaging operator, which is defined by

$$a_{ij} = \frac{1}{K}\left[a_{ij}^1 \ (+) a_{ij}^2 \ (+) \ldots\ldots\ldots + a_{ij}^j\right] \ ,$$

Where K is the number of experts and a_i^j (6)

Is the evaluation of the K[th] decision maker on the pair-wise importance comparison of j[th] and k[th] criteria.

Step IV: Obtain the fuzzy weights W_i. The derivation of $\tilde{z}_j$ values (Eq. (2)) and fuzzy weights $\tilde{W}j$ (Eq. (3)) can be detailed as follows.

Let:

$$a_j = \left[\prod_{k=1}^{n} a_{jk}\right]^{1/n} \text{ and } a_j = \sum_{j=1}^{n} a_j \text{ ----------(7)}$$

The fuzzy weight ($\tilde{w}_i$) of every attribute is computed in Eq. (3) which involves 3 sub-processes.Initially, the vector summation of every $\tilde{r}_i$ is found. Next, the inverse of summation vector is done andreplaces the T_rFN. Finally, the $\tilde{W}_i$ is calculated by multiplying every $\tilde{r}_i$ with its reverse vector.

Step V: Defuzzify and normalize the triangulalr fuzzy weights.

To defuzzify the TrFN in Eq. (8), the following equation is used:

$$W_i = \frac{l_1 + m_1 + u_1}{3} \quad (8)$$

And, to normalize the crisp weights Eq. (9) is used: $W_j = \frac{w_j'}{\sum_{j=1}^{n} w_j'}$, $j=1,2,.....n$ (9)

MODEL VALIDATION: A CASE OF INDIAN FMCGS COMPANY

Fast Moving Consumer Goods (FMCGs) industries captures oligopolistic market, so being sustain in the global market, recurrently measurement of performance become essential task. In order to determine the performance of the implemented total quality management which strongly depends on its critical success factors and capabilities, the linguistic data is collected including all levels of management viz, Low level, Medium level and Top level through questionnaire. The involvement of all three levels is to nullify the biasness of the response provided and meet the requirements of TQM as well and observed that mostly using chomskyan linguistics. The vision behind involvement the employees of three levels to form a quality circle. As it is well known fact that the quality circle helps to find out relationship between different factors in the organization and their effect on each other so that we can find out major and minor factors based upon ranking which management can take action to improve organization effectiveness. Zbaracki (1998) the concept of TQM provides a simple and essential description of quality management that allows middle and upper level managers to oversee product and service quality in direct and uncomplicated manner. The primary objective associated with modeling for performance measurement is to identify the various critical factors, interacting operations in a realistic manner. Achieving this objective is complicated by existing TQM barriers, if any. Given the complexity of operations and the uncertainty associated with several of these factors, FAHP was identified as an important decision-making tool for use in understanding, quantifying, and improving performance. During an actual study of organization and its operation environment, a decision maker will monitor multiple performance measures. These include existing TQM drivers, enablers, and capabilities for all types of factors. TQM drivers represent the bottom line and are considered the most important performance measure. TQM enablers, which is positively correlated with drivers, and both reinforce the capability of TQM. In total 14 attributes were considered for 03 criteria for the present condition (Table 1). On the basis of collected information hierarchy is developed for the evaluation (Figure 1).

Table 1: Performance Measurement Criteria of TQM

Level I	Level II	Level III	Level IV
GOAL	CRITERIA	ALTERNATIVES	
TQM Objectives/Goal	TQM Drivers	Top Level Management	Organizational Performance
		Customer Focus	
		Continuous improvement and innovation	
		Supplier management	
		Employee involvement	
		Quality Information /Information Quality	
		Benchmarking	
	TQM Enablers	Culture and communication	
		Employee Encouragement and Motivation	
		Training and education	
	TQM Capabilities	Zero Defect	
		Waste Reduction	
		Continuous Improvement Process	
		Performance Improvement	

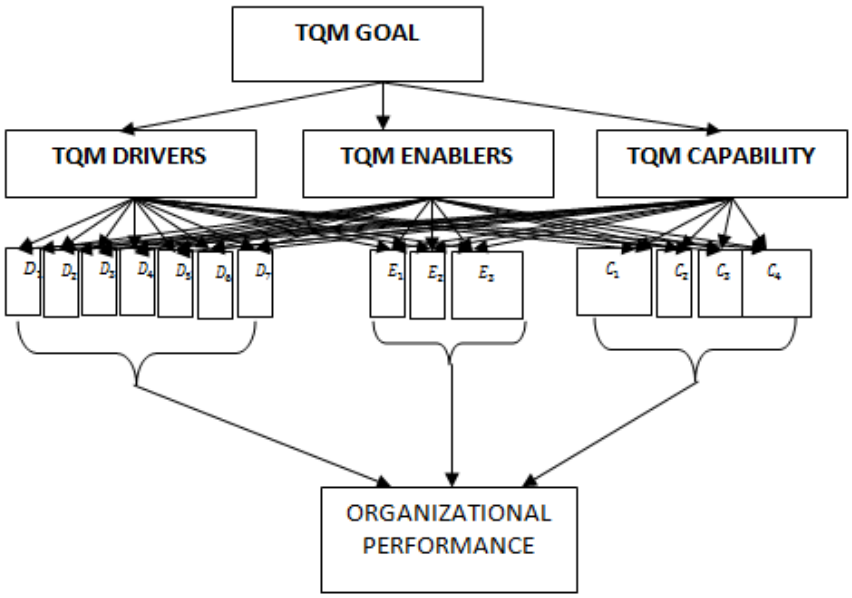


Figure 1 Hierarchy Based Performance Measurement Model

TQM Drivers: D_1-Top Level Management; D_2-Customer Focus/Customer Centricity;D_3- Continuous improvement and innovation;D_4- Supplier management;D_5-Employee involvement;D_6-Quality Information /Information Quality;D_7-Benchmarking

TQM Enablers: E_1- Culture and communication;E_2- Employee encouragement and motivation;E_3- Training and education

TQM Capability: C_1-Zero Defect;C_2-Waste Reduction;C_3- Continuous Improvement Process;C_4-Performance Improvement

To implement the method, it was necessary the support of an expert in decision-making methods and of those engaged for TQM implementation in the FMCGs industry studied here. The support of the decision-making expert was necessary since he was responsible for telling the TQM staff how the judgments should be performed, how the program should be used and what maximum value should the inconsistency take. Experienced people were chosen from and formed three groups of evaluators. Table 2 presents the comparisons among TQM practices, provided by Group experts. The comparison matrix presented in Table 2 is consistent. Remarkably, all the comparison matrices provided by the Group experts were consistent. These are indications that the experts were aware about their judgement.

Table 2: Pair-Wise Comparison of the Three Main Criteria with Respect to the TQM Practices

TQM Goal based on	DRIVERS	ENABLERS	CAPABILLITIES	PRIORITY
DRIVERS	1	3	5	65.2 %
ENABLERS	1/3	1	2	22.3 %
CAPABILLITIES	1/5	1/2	1	12.5 %

The Aggregation of Individual Priorities (Saaty and Peniwati, 2007) was applied to aggregate their comparisons. The experts were considered as equally important. That is, their priorities had the same weight when aggregated.

Explanation of the attributes of TQM used is as:

Oakland (1993) suggested ten basic steps in the implementation of TQM projects from that the extracted points which emphasizes critical success factors are as about: Develop the mission onto its critical success factors (CSFs) to coerce and move it forward, Breakdown the critical success factors into the key or critical process and gain process ownership as-well-as prepare project quality plans for all levels of work and Breakdown the critical success factors into sub-processes, activities and tasks and form improvement teams around these.

Top level management (D_1) Top level management is the one, which is responsible for a firm's strategy, can also be judged how performance is scaled for improvement. The majority of respondents reported "the lack of complete top management commitment from the start, and through the process" to be the obvious barrier to TQM success. The achievement of organizational goals is based on top management teams' ability to anticipate and respond to external change (Burgelman, 1991; Child, 1972).W. C. Auden et al. (2006) emphasized as summary of the work that the top management team seems to have a strong influence on the firm's performance.

Customer focus/Customer Centricity(D_2) Customer Focus involves researching what customers or end users want from products or services. Prajogo and Cooper (2010) argue that quality customer service is one of the most important aspects of TQM because in any customer-oriented business the customers are the organisation's main stakeholders. Nair (2006) highlighted the importance of customer focus by demonstrating evidence of a strong relationship between customer focus and firm performance. Moreover, customer focus receives attention as a CSF, because every effort of TQM is a customer orientated.

*Continuous improvement and innovation (D_3)*TQM practices enhance competitive performance by improving the quality of products and processes. The continuous improvement leads to dynamic outcomes of innovation, namely product innovation – changes in the specific products/services offered to the customers and process innovation – and changes in the mode by which the products are created or delivered. Innovation refers to an outcome perceived as new, regardless of whether it is an idea, object or process, as well as to the process of creating this newness (Slappendel, 1996).

Supplier's management (D_4) supplier's management acts as an interface between TQM and Supply Chain Management (SCM). Managing the flow of goods between the manufacturer and customers is a fundamental part of organizations. To accomplish this the well-trained employees, ensure consistent, reliable and secure supply. Material traceability is always also maintained through robust quality management systems.Vanichchinchai, A. and Igel, B. (2011) investigated the relationships among total quality management practices (TQMP), supply chain management practices (SCMP) and firm's supply performance (FSP) in the automotive industry in Thailand and found that the set of SCMP, TQMP and FSP measures are reliable and valid. Gunasekaran and McGaughey (2003) earlier also suggested that TQM could play a key role in improving SCM.

Employee Involvement: *(D_5)* The negative implications of TQM and performance management inconsistency is also compounded by the fact that employee's engagement in designing and implementing the performance management system, in the majority of organizations, is quite low. The survey on employee engagement reported in Gallup Management Journal (2002) according to the report (see, for further details, gmj.gallup.com), "more than half of the employees may not be engaged with their work". Involvement of employee automatically empowers the people to make decisions, make choices, take risks, and make new contributions. This congruity, in turn, will improve employee motivation and attitudes towards effective implementation of TQM initiatives.

Quality Information/Information Quality (D_6) To have an effective TQM implementation, organization must ensure that data are valid and reliable. They measure what they are supposed to consistently, and employees should have access to the data that are needed to carry out the tasks, Seetharaman, A. et al. (2006). Lakhal et al. (2006) establish through study and confirm the relevance of data-based factual decision-making by pointing out that information and analysis have a significant direct effect on various performance measures. Information and analysis and plant performance are correlated positively, years back shown by Choi and Eboch (1998) in their study.

*Benchmarking (D_7)*Benchmarking aims to measure organization's operations or processes against the best-in-class performers from inside or outside its industry. The usefulness of benchmarking for improving the performance of the organization and to achieve competitive advantagehelp in continuous service improvements and establishment of customer satisfaction. Benchmarking can significantly affect the improvement of key business processes and consequently, increase the quality level, Hietschold, N. et al. (2014). Without benchmarking, organisations do not know their relative performance and they probably fail to design processes more effectively (Rao, S. et al. 1999).

Culture and communication (E_1): Competent organizational culture can be defined as, a set of congruent behaviors, attitudes, and policies that come together in a system, agency, or among professionals that enable them to work effectively in cross cultural situations.(Saldana, 2001). Culture is important to improve the communication among top-to – bottom, bottom-to-top and across the departments, in which the information is shared by each staff. Antony, J. et al. (2002). As a concept, 'Quality' implies that an organization understands what customers require of a product or service, and is able not only to meet this requirement but can also exceed the expectations of the customer. This demands a particular culture that can in fact, be seen to be embedded in TQM philosophy, which embraces the idea of encouraging an organization to strive for continuous improvement in its values and methods of operation in order to achieve customer satisfaction.

Employee encouragement and motivation(E_2) White and Nebeker (1996, p. 210) in their review of "reward and recognition and their functions in TQM organizations" point out that: A team-based perspective that focus on process improvement and personal development could improve effectiveness of performance appraisal as well as TQM objectives. Soltani, E. el al. (2005) survey indicated that their performance management systems in relation to having a positive impact on employee motivation towards successful implementation of quality programs were "very" and "moderately" effective.

Training and education ($\boldsymbol{E_3}$) Mital et al. (1999) along with group of fifteenco-authors recommend that there is a dire need to train workers in manufacturing organizations and thereby improve the overall effectiveness and efficiency of such organizations. Further also mentioned that in order to compete successfully in the global market, manufacturing organizations must aim at training workers in skills necessary to produce quality goods. Strong correlation between quality performance and employee training is established in a study by Solis et al. (2000).

Zero-defect ($\boldsymbol{C_1}$) Zero Defects focuses on conditioning the employee to ensure a personal interest serves as motivation in everything they do. This is done by motivating the employee towards work importantly for the success of the organization. Specific goals, based on the customer's requirements (Linderman, K. et al., 2003) that are grounded on a philosophy of continuous improvements, all the points that combines TQM, Zero Defect Mindset (ZDM) and Six Sigma together. Total quality management models tend to incorporate the zero-defect vision by maximizing the overlap of customer demands and delivered product features, while costs have to be minimized, M. Colledani et al.(2014).

Waste reduction (C_2) The increase in lead-time also causes quality related phenomena which reduces the observability of the critical product quality feature in the system, further reduces the ability to detect a potentially generated defect within a short time. Colledani et al. (2014) stated that this directly translates into the propagation of more defects between the processing stages in the system and a waste of production capacity in processing parts that are already defective. A loss of capacity leads to a loss of total throughput.

*Continuous improvement process (C_3)*Walsh, A. et al. (2002), outlook that many organizations show great enthusiasm for TQM during the initial year of TQM initiatives but then fall victim to a lackluster attitude towards its continuation. Lastly authors concluded with emphasizing that it is most important to remember that successful TQM strategies are based around the concept of continuous improvement. The continuous improvement of the process leads to customer satisfaction, which results in an external quality improvement. The continuous improvement also leads to fewer defects, which results in an internal quality improvement (Dahlgaard et al., 1994).

Performance improvement (C_4) TQM is not limited to shop floor; it may be applied throughout the enterprise. This encompasses whole functions of the organization. TQM provides a single, unifying approach for the entire enterprise using performance improvement as the capability. This capability of TQM is important for organizations that offer high-variety and custom products today, but it will become increasingly important as we look to the future with mass customization.

To achieve the goal of the TQM through existing drivers, enablers and capabilities is shown in figure 1 with respective factors as sub-criteria. This work may be extended by considering the sub sub-criteria of each factor, which is not considered in this research.

Performance Measurement Based on Fuzzy AHP Methodology

After determining the evaluation criteria and the alternatives, the steps of the integrated fuzzy AHP algorithm are implemented. To obtain the importance of each factor first the experts utilised saaty's nine point scale, and result is obtained. Table 3 gives the results of the pair-wise comparisons of the evaluation criteria made by experts. Tables 2 and 3, the aggregated fuzzy evaluation matrix for the criteria weights is obtained as in Table 4.

Table 3: Pair-Wise Comparison of Attributes of TQM

TQM		Drivers							Enablers			Capabilities			
		D_1	D_2	D_3	D_4	D_5	D_6	D_7	E_1	E_2	E_3	C_1	C_2	C_3	C_4
Drivers	D_1	**1**	9	5	7	3	7	5	4	4	6	4	5	5	7
	D_2	1/9	**1**	2	5	2	4	2	2	2	2	4	2	2	3
	D_3	1/5	1/2	**1**	2	3	7	4	7	5	3	8	8	9	5
	D_4	1/7	1/5	1/2	**1**	2	8	2	3	3	2	6	5	4	5
	D_5	1/3	½	1/3	1/2	**1**	5	4	7	9	5	7	7	6	6
	D_6	1/7	1/4	1/7	1/8	1/5	**1**	3	4	3	3	5	4	5	7
	D_7	1/5	1/2	1/4	1/2	1/4	1/3	**1**	3	3	2	4	4	4	6
Enablers	E_1	1/4	1/2	1/7	1/3	1/7	1/4	1/3	**1**	5	3	5	5	6	4
	E_2	1/4	1/2	1/5	1/3	1/9	1/3	1/3	1/5	**1**	3	5	5	7	5
	E_3	1/6	1/2	1/3	1/2	1/5	1/3	1/2	1/3	1/3	**1**	3	3	5	5
Capabilities	C_1	1/4	1/4	1/8	1/6	1/7	1/5	1/4	1/5	1/5	1/3	**1**	9	7	9
	C_2	1/5	1/2	1/8	1/5	1/7	1/4	1/4	1/5	1/5	1/3	1/9	**1**	5	7
	C_3	1/7	1/3	1/5	1/5	1/6	1/7	1/6	1/4	1/5	1/5	1/9	1/7	**1**	9
	C_4	1/7	1/3	1/5	1/5	1/6	1/7	1/6	1/4	1/5	1/5	1/9	1/7	1/9	**1**

For pairwise comparisons of criteria and alternatives, Saaty (1994) defined a scale of integers from 1 to 9 that express how many times more important or dominant one attribute is over another one. To each number from the scale assigned a linguistic term describing the intensity of importance or dominance of one attribute over another. Generally, integers 1, 3, 5, 7 and 9 are used for making pair wise comparisons of attributes. In case where more detailed comparisons are desirable then intermediate values 2, 4, 6, and 8 is selected. A square pair wise comparison matrix $A = [a_{ij}]_{ij=1}^{p}$ is constructed ($p = 14$, in our case), Table 3. An element a_{ij} expresses how many times more important $x_i$is over $x_j$with respect to a particular object from the upper level of the hierarchy. It is obvious that a_{ii}= 1, i= 1, . . . , p, as we compare x_i with itself. Further, when object x_i is a_{ij}-times more important than object x_j, then obviously object x_j takes only the $1/a_{ij}$-th of the importance of object x_i. Therefore, the pairwise comparison matrix $A = [a_{ij}]_{ij=1}^{p}$ is reciprocal, i.e. a_{ji}= $1/a_{ij}$for all $i,j \in \{1, \ldots, p\}$.

As this work is carried in an Indian FMCGs industry, so in real-world decision- making problems, the information provided by decision makers is often vague because the information is given by means of linguistically. Thus, fuzzy AHP has been developed in order to deal with vagueness and imprecision. Many different approaches to the fuzzification of Saaty's scale have been proposed in the literature. The authors expressed the meaning of the linguistic terms by triangular fuzzy numbers.

Using the elements from a considered fuzzified scale, a fuzzy pairwise comparison matrix $A = [a_{ij}]_{ij=1}^{p}$ is constructed (Table 4), i.e. $\tilde{a}_{ij} = 1\tilde{}$, i= 1, . . . , p, and $\tilde{a}_{ji}$= $1/\tilde{a}_{ij}$for all $i,j \in \{1, \ldots, p\}$.

Table 4: With Fuzzy Numbers

TQM		Drivers							Enablers			Capabilities			
		D_1	D_2	D_3	D_4	D_5	D_6	D_7	E_1	E_2	E_3	C_1	C_2	C_3	C_4
Drivers	D_1	**1,1,1**	9,9,9	4,5,6	6,7,8	2,3,4	6,7,8	4,5,6	3,4,5	3,4,5	5,6,7	3,4,5	4,5,6	4,5,6	6,7,8
	D_2	1/9,1/9,1/9	**1,1,1**	1,2,3	4,5,6	1,2,3	3,4,5	1,2,3	1,2,3	1,2,3	1,2,3	3,4,5	1,2,3	1,2,3	2,3,4
	D_3	1/6,1/5,1/4	1/3,1/2,1	**1,1,1**	1,2,3	2,3,4	6,7,8	3,4,5	6,7,8	4,5,6	2,3,4	7,8,9	7,8,9	9,9,9	4,5,6
	D_4	1/8,1/7,1/6	1/6,1/5,1/4	1/3,1/2,1	**1,1,1**	1,2,3	7,8,9	1,2,3	2,3,4	2,3,4	1,2,3	5,6,7	4,5,6	3,4,5	4,5,6
	D_5	1/4,1/3,1/2	1/3,1/2,1	1/4,1/3,1/2	1/3,1/2,1	**1,1,1**	4,5,6	3,4,5	6,7,8	9,9,9	4,5,6	6,7,8	6,7,8	5,6,7	5,6,7
	D_6	1/8,1/7,1/6	1/5,1/4,3/4	1/8,1/7,1/6	1/9,1/8,1/7	1/6,1/5,1/4	**1,1,1**	2,3,4	3,4,5	2,3,4	2,3,4	4,5,6	3,4,5	4,5,6	6,7,8
	D_7	1/6,1/5,1/4	1/3,1/2,1	1/5,1/4,3/4	1/3,1/2,1	1/5,1/4,3/4	1/4,1/3,1/2	**1,1,1**	2,3,4	2,3,4	1,2,3	3,4,5	3,4,5	3,4,5	5,6,7
Enablers	E_1	1/5,1/4,3/4	1/3,1/2,1	1/8,1/7,1/6	1/4,1/3,1/2	1/8,1/7,1/6	1/5,1/4,1/3	1/4,1/3,1/2	**1,1,1**	4,5,6	2,3,4	4,5,6	4,5,6	5,6,7	3,4,5
	E_2	1/5,1/4,3/4	1/3,1/2,1	1/6,1/5,1/4	1/4,1/3,1/2	1/9,1/9,1/9	1/4,1/3,1/2	1/4,1/3,1/2	1/6,1/5,1/4	**1,1,1**	2,3,4	4,5,6	4,5,6	6,7,8	4,5,6
	E_3	1/7,1/6,1/5	1/3,1/2,1	1/4,1/3,1/2	1/3,1/2,1	1/6,1/5,1/4	1/4,1/3,1/2	1/3,1/2,1	1/4,1/3,1/2	1/4,1/3,1/2	**1,1,1**	2,3,4	2,3,4	4,5,6	4,5,6

Capabilities	**C 1**	1/5,1/4,3/4	1/5,1/4,3/4	1/9,1/8,1/7	1/7,1/6,1/5	1/8,1/7,1/6	1/6,1/5,1/4	1/5,1/4,3/4	1/6,1/5,1/4	1/6,1/5,1/4	1/4,1/3,1/2	**1,1,1**	9,9,9	6,7,8	9,9,9
	C 2	1/6,1/5,1/4	1/3,1/2,1	1/9,1/8,1/7	1/6,1/5,1/4	1/8,1/7,1/6	1/5,1/4,3/4	1/5,1/4,3/4	1/6,1/5,1/4	1/6,1/5,1/4	1/4,1/3,1/2	1/9,1/9,1/9	**1,1,1**	4,5,6	6,7,8
	C 3	1/8,1/7,1/6	1/4,1/3,1/2	1/6,1/5,1/4	1/6,1/5,1/4	1/7,1/6,1/5	1/8,1/7,1/6	1/7,1/6,1/5	1/5,1/4,3/4	1/6,1/5,1/4	1/6,1/5,1/4	1/9,1/9,1/9	1/8,1/7,1/6	**1,1,1**	9,9,9
	C 4	1/8,1/7,1/6	1/4,1/3,1/2	1/6,1/5,1/4	1/6,1/5,1/4	1/7,1/6,1/5	1/8,1/7,1/6	1/7,1/6,1/5	1/5,1/4,3/4	1/6,1/5,1/4	1/6,1/5,1/4	1/9,1/9,1/9	1/8,1/7,1/6	1/9,1/9,1/9	**1,1,1**

For $i = j$, there is obviously no fuzziness in the comparison, and thus, $\tilde{a}_{ji}$ has to be set as $(1, 1, 1)$ for any $i \in \{1, \ldots, p\}$. On the other hand, when two different attributes x_i and x_j, $i = j$, are assessed to be "equally important", then this comparison may contain some fuzziness.

K Teknomo,(2006) suggested that in practice, don't need measurement of inconsistency Random Index (RI) for more than n=7 because the number of pair comparisons n*(n-1)/2 become so large. Too many questions on the same things make people confuse to answer them correctly. Utilize the hierarchy by grouping some criteria into a higher level hierarchy of the criteria. As in this present study number of attributes considered by authors is 14, which is more than 7, so random indexing is exempted.

When the pairwise comparison matrix of objects is constructed, weights of the factors can be computed from the matrix. Based on Buckley's (1985) method, the geometric mean (Table 5) of fuzzy comparison value of every attribute in computed and formulated in equation below. This method is used because it is easy to extend to the fuzzy case and guarantees a unique solution to the reciprocal comparison matrix. The weight assessing method by geometric mean is chosen for its simplicity and ease in its application to the fuzzy case, Kaya, T., and Kahraman, C. (2011).

Formula used to obtain Fuzzy Geometric Mean Value of each attribute is calculated as

$$\tilde{A}_1 \otimes \tilde{A}_2 = (\ell_1, m_1, u_1) \otimes (\ell_2, m_2, u_2)$$

$$= (\ell_1 + \ell_2, m_1 + m_2, u_1 + u_2)$$

Table 5: Fuzzy Geometric Mean

	Fuzzy Geometric Mean values			
	l	**m**	**u**	
D_1	3.807	4.652	5.460	3.807, 4.652,5.460
D_2	1.160	1.879	2.529	1.160,1.879,2.529
D_3	2.363	3.013	3.730	2.363,3.013,3.730
D_4	1.316	1.836	2.395	1.316,1.836,2.395
D_5	1.998	2.467	3.166	1.998,2.467,3.166
D_6	0.937	1.176	1.432	0.937,1.176,1.432
D_7	0.848	1.151	1.576	0.848,1.151,1.576
E_1	0.768	0.968	1.252	0.768,0.968,1.252
E_2	0.651	0.812	1.051	0.651,0.8112,1.051
E_3	0.548	0.727	1.043	0.548,0.727,1.043
C_1	0.436	0.505	0.606	0.436,0.505,0.606
C_2	0.315	0.377	0.472	0.315,0.377,0.472
C_3	0.235	0.270	0.320	0.235,0.270,0.320
C_4	0.171	0.197	0.234	0.171,0.197,0.234

Then Fuzzy weight is calculated of each attribute using formula

$\tilde{W}_i = \tilde{r}_1 \otimes (\tilde{r}_1 \oplus \tilde{r}_2 \ldots\ldots\ldots\ldots \oplus \tilde{r}_n)^{-1}$ Also, $\tilde{A}_1 \oplus \tilde{A}_2 = (\ell_1, m_1, u_1) \oplus (\ell_2, m_2, u_2)$

$= (\ell_1 + \ell_2, m_1 + m_2, u_1 + u_2)$

Table 6: Fuzzy Weights

	Fuzzy Weights			
	l	**m**	**u**	
D_1	0.151	0.232	0.351	0.151,0.232,0.351
D_2	0.046	0.094	0.163	0.046,0.094,0.163
D_3	0.094	0.150	0.240	0.094,0.150,0.240
D_4	0.052	0.092	0.154	0.052,0.092,0.154
D_5	0.079	0.123	0.204	0.079,0.123,0.204
D_6	0.037	0.059	0.092	0.037,0.059,0.092
D_7	0.034	0.057	0.101	0.034,0.057,0.101
E_1	0.030	0.048	0.081	0.030,0.048,0.081
E_2	0.026	0.041	0.068	0.026,0.041,0.068
E_3	0.022	0.036	0.067	0.022,0.036,0.067
C_1	0.017	0.025	0.039	0.017,0.025,0.039
C_2	0.012	0.019	0.030	0.012,0.019,0.030
C_3	0.009	0.013	0.021	0.009,0.013,0.021
C_4	0.007	0.010	0.015	0.007,0.010,0.015

Fuzzy weights are defuzzified using Centre of area (COA) method to get the crisp values for all the factors considered for the study and then normalised. The value obtained $w_{total} > 1$, so it must be normalised. On the basis of the normalised values the ranking of TQM attributes demonstrated that impact on its performance. It is observed that Top level management (D_1) is given the highest ranking for these factors with normalised weight (0.180). The rankings of collected factors for each category aggregately are computed in a similar manner.

Table 7: Priority Weights of Alternatives

			De-fuzzified weights using centroid method $W_i = \frac{a_i + m_i + b_i}{3}$	normalised weights $\frac{W_i}{\sum_{i=1}^{n} W_i}$	Ranking
	D_1	0.151,0.232,0.351	0.244646176	0.180	1
	D_2	0.046,0.094,0.163	0.100766316	0.079	4
	D_3	0.094,0.150,0.240	0.161247026	0.120	2
	D_4	0.052,0.092,0.154	0.099234373	0.076	5
	D_5	0.079,0.123,0.204	0.135261797	0.101	3
	D_6	0.037,0.059,0.092	0.062610712	0.047	7
	D_7	0.034,0.057,0.101	0.064129672	0.049	6
	E_1	0.030,0.048,0.081	0.053090707	0.040	8
	E_2	0.026,0.041,0.068	0.044620883	0.033	9
	E_3	0.022,0.036,0.067	0.041682472	0.032	10
	C_1	0.017,0.025,0.039	0.027134575	0.020	11
	C_2	0.012,0.019,0.030	0.020549803	0.015	12
	C_3	0.009,0.013,0.021	0.014453461	0.011	13
	C_4	0.007,0.010,0.015	0.010559698	0.008	14
Total of De-fuzzified weights, w_{total}			1.079987671		

For interpretation of the result the rank is considered and termed them that they are acting as barriers to the TQM and must be take care of. As Performance improvement (C_4) is measured 14th rank, is the most potential barrier identified.

In this way, the proposed performance model/framework is for the implemented TQM process, is evolvable, perfective or preventive changes may be introduced when needed, is specific for target FMCGs, more CSFs have been involved at different levels and has been evaluated mathematically, and then validated appropriately.

MANAGERIAL IMPLICATION AND CONCLUSION

In this practical work integrated FAH approach is proposed for solving a complex decision-making task related to TQM. To validate the potentiality of our approach we have analysed the implemented TQM of FMCGs industry. In the published research work, FAHP, based on the extent analysis method was used to derive the relative importance between the critical success factors and organisational performance. We have used the same pairwise comparison matrix as given in this original work and have utilized Buckley's FAHP approach (geometric mean approach) to calculate the relative importance among the risk factors. In the next task, to rank the attributes, we have proposed the normalization formulae to calculate the rank in practical manner. We believe that our approach is computationally faster, more accurate and requires less computational steps than the other performance assessment method. Although we have not compared our ranking result with any other work of the FAHP. Apart from that, when the number of critical success factors and alternatives are increased, our approach is sufficient to handle that situation and still provides credible ranking results. Next, we have proposed a quite simpler to interpret and is derived from simpler human reasoning. This integrated approach is believed to be a feasible choice among the performance measurement in industries to support the performance assessment task of other incorporated sub-systems like SCM, Six-Sigma, JIT, TPM etc. Apart from that, other weight calculation approaches may be use to improve the decision-making process.

Funding: This research did not receive any specific grant from agencies in the public, commercial, or not-for-profit sectors.

Declarations of Interest: None

REFERENCES

1. *Ahire, S. L., Waller, M. A. & Golhar, D. Y. (1996). Quality management in TQM versus non-TQM firms: an empirical investigation. International Journal of Quality 13(8): 8–27.*
2. *Auden, W. C., Shackman, J. D., & Onken, M. H. (2006). Top management team, international risk management factor and firm performance. Team Performance Management: An International Journal, 12(7/8), 209–224. doi:10.1108/13527590610711778*
3. *Anu P. Anil, Dr. Satish K. P. (2015), "Investigating the relationship between TQM practices and Firm's performance: A conceptual framework for Indian organizations", International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST- 2015), Procedia Technology 24 (2016) 554 – 561, doi: 10.1016/j.protcy.2016.05.103*
4. *Antony, J., Leung, K., Knowles, G., & Gosh, S. (2002). Critical success factors of TQM implementation in Hong Kong industries. International Journal of Quality & Reliability Management, 19(5), 551–566. doi:10.1108/02656710210427520*
5. *Barot, Digvijay., et al. "An empirical assessment of smart phone usage amongst students and professors." International Journal of Business and General Management (IJBGM) 3 (2014): 19-28.*

6. *Buckley, J. J. (1985) extended Saaty's AHP method to incorporate fuzzy comparison ratios aij and uses the geometric mean method to derive fuzzy weights and performance scores.*
7. *Buckley, J. J. (1985). Fuzzy hierarchical analysis. Fuzzy Sets and Systems, 17(2), 233-247. https://doi.org/10.1016/0165-0114(85)90090-9*
8. *Bellman, R. E., & Zadeh, L. A. (1970). Decision-making in a fuzzy environment. Management Science, 17(4), 141–164.*
9. *Bratton, J. and J. Gold (1999). Human Resource Management: Theory and Practice (2nd edn). Macmillan Press, London.*
10. *Burgelman, R. (1991), "Intra-organizational ecology of strategy making and organizational adaptation: theory and field research", Organization Science, Vol. 2, pp. 239-62.*
11. *Child, J. (1972), "Organization structure, environment, and performance: the role of strategic choice", Journal of Management Studies, Vol. 11, pp. 13-27.*
12. *Cheng, E.W.L and Li,H. (2002), "Analytic hierarchy process: a decision making method for construction management", Proceedings of the first International conference on construction in the 21st century (CITC2002), Challenges and opportunities in the Management and Technology, 25-26 April, Miami,FL*
13. *Chou TY, Hsu CL, Chen MC (2008) A fuzzy multi-criteria decision model for international tourist hotels location selection. Int J Hosp Manag 27(2):293–301*
14. *Choi, Thomas Y., and Karen Eboch. 1998. "The TQM Paradox: Relations among TQM Practices, Plant Performance, and Customer Satisfaction." Journal of Operations Management 17 (1): 59–75.*
15. *Colledani, M., Tolio, T., Fischer, A., Iung, B., Lanza, G., Schmitt, R., & Váncza, J. (2014). Design and management of manufacturing systems for production quality. CIRP Annals, 63(2), 773–796. doi:10.1016/j.cirp.2014.05.002*
16. *Dessler, G. (2000). Human resource management (8th ed.). New Jersey: Prentice-Hall.*
17. *Dess, G. G., & Robinson, R. B. (1984). Measuring organizational performance in the absence of objective measures: The case of the privately-held firm and conglomerate business unit. Strategic Management Journal, 5(3), 265–273. doi:10.1002/smj.4250050306*
18. *Demirbag, M., Tatoglu, E., Tekinkus, M. and Zaim, S. (2006), "An analysis of the relationship between TQM implementation and organizational performance: evidence from Turkish SMEs", Journal of Manufacturing Technology Management, Vol. 17 No. 6, pp. 829-47.*
19. *Dahlgaard, J.J., Kristensen, K. and Kanji, G.K. (1994), The Quality Journey: A Journey Without an End, Carfax, Arbingdon.*
20. *Eisenhardt, K.M., Kahwajy, J. and Bourgeois, L. (1998), "Conflict and strategic choice: how top management teams disagree", in Hambrick, D., Nadler, D. and Tushman, M. (Eds), Navigating Change: How CEOs, Top Teams, and Boards Steer Transformation, Harvard Business School Press, Boston, MA.*
21. *Feigenbaum, A.V. (1988) Total quality control: achieving productivity, market penetration and advantage in the global economy, McGraw-Hill Higher Education.*
22. *Gallup Management Journal (2002), "Start worrying about "not engaged" employees", Gallup Management Journal, available at: http://gmi.gallup.com.*
23. *Gunasekaran, A. and McGaughey, R.E., 2003. TQM in supply chain management. The TQM Magazine, 15 (6), 361–363.*

24. Huang, L-C and Wu, R. Y-H. Applying fuzzy analytic hierarchy process in the managerial talent assessment model – an empirical study in Taiwan's semiconductor industry. Intern. Journal of Technology Management, Vol 30, No. ½. 2005. pp 105-130.

25. Hietschold, N., Reinhardt, R., & Gurtner, S. (2014). Measuring critical success factors of TQM implementation successfully – a systematic literature review. International Journal of Production

26. Ismail Salaheldin, S. (2009). Critical success factors for TQM implementation and their impact on performance of SMEs. International Journal of Productivity and Performance Management, 58(3), 215–237. doi:10.1108/17410400910938832

27. Jiansheng Zhangab and Wei Tana (2012)," Research on the Performance Evaluation of Logistics Enterprise Based on the Analytic Hierarchy Process", International Conference on Advances in Energy Engineering, Energy Procedia 14 (2012) 1618 – 1623

28. Kayakutlu G, Buyukozkan G (2011) Assessing performance factors for a 3PL in a value chain. International Journal of Production Economics, 131(2):441–452

29. Koh, S., Demirbag, M., Bayraktar, E., Tatoglu, E. and Zaim, S. (2007), "The impact of supply chain management practices on performance of SMEs", Industrial Management & Data Systems, Vol. 107 No. 1, pp. 103-24.

30. K Teknomo, "Analytic hierarchy process (AHP) tutorial", Revoledu. com, 2006

31. Kaya, T., &Kahraman, C. (2011). An integrated fuzzy AHP–ELECTRE methodology for environmental impact assessment. Expert Systems with Applications, 38(7), 8553–8562.

32. Kabir, G.; Akhtar Hasin, M.A. Multi-criteria inventory classification through integration of fuzzy analytic hierarchy process and artificial neural network. Int. J. Ind. Syst. Eng. 2013, 14, 74–103.

33. Kumar, P., & Singh, R. K. (2012). A fuzzy AHP and TOPSIS methodology to evaluate 3PL in a supply chain. Journal of Modelling in Management, 7(3), 287–303. doi:10.1108/17465661211283287

34. Liou, James J.H., Tzeng, G.H., 2007. A non-additive model for evaluating airline service quality. Journal of Air Transport Management 13 (3), 131–138.

35. Li, Wenhua & Yu, Suihuai & Pei, Huining & Zhao, Chuan & Tian, Baozhen, 2017. "A hybrid approach based on fuzzy AHP and 2-tuple fuzzy linguistic method for evaluation in-flight service quality," Journal of Air Transport Management, Elsevier, vol. 60(C), pages 49-64. DOI: 10.1016/j.jairtraman.2017.01.006

36. Linderman, K., Schroeder, R. G., Zaheer, S. & Choo, A. S., 2003. Six Sigma: a goal-theoretic perspective. Journal of Operational Management 21(2003), 21(2003), pp. 193-203.

37. Lakhal, Lassâad, Federico Pasin, and Mohamed Limam. 2006. "Quality Management Practices and their Impact on Performance." International Journal of Quality & Reliability Management 23 (6): 625–646.

38. Motwani, J.G., Mahmoud, E. and Rice, G. (1994), "Quality practices of Indian organizations:

39. an empirical analysis", International Journal of Quality & Reliability Management, Vol. 11 No. 1, pp. 38-52.

40. Moon, P. (1993). Appraising Your Staff. Kogan Page, London.

41. Mital, A., Pennathur, A., Huston, R.L., Thompson, D., Pittman, M., Markle, G., Kaber, D.B.,Crumpton, L., Bishu, R.R., Rajurkar, K.P., Rajan, V., Fernandez, J.E., McMulkin, M., Deivanayagam, S., Ray, P.S. and Sule, D. (1999),The need for worker training in advanced manufacturing technology (AMT) environments: a white paper, International Journal of Industrial Ergonomics, Vol. 24, No. 2, pp.173–184.

42. *Nair, Anand. 2006. "Meta-analysis of the Relationship between Quality Management Practices and Firm Performance – Implications for Quality Management Theory Development." Journal of Operations Management 24 (6): 948–975.*
43. *Oakland, J. S. (1993). Implementation of TQM and Management of Change, Total Quality Management: The Route to Improving Performance Jordan Hill, Oxford: Butterworth Heinemann Ltd.*
44. *Odukoya, Jonathan Adedayo, et al. "Employers'perception of Covenant University Graduates." IASET: International Journal of Library and Educational Science 2.3 (2016).*
45. *Prajogo, D.I. &Sohal, A.S. (2006) 'The relationship between organization strategy, total quality management (TQM), and organization performance—the mediating role of TQM', European Journal of Operational Research, vol. 168, no. 1,pp. 35-50.*
46. *Prajogo, D.I. and Sohal, A.S. (2004), "The multi dimensionality of TQM practices in determining quality and innovation performance – an empirical examination", Technovation, Vol. 24 No. 6, pp. 443-53.*
47. *Pfeffer, J. & Sutton, R.I. (2006) 'Evidence-based management', Harvard business review, vol. 84, no. 1,p. 62*
48. *Prajogo, D.I. & Cooper, B.K. (2010) 'The effect of people-related TQM practices on job satisfaction: a hierarchical model', Production Planning and Control, vol. 21, no. 1,pp. 26-35*
49. *Rakesh Kumar, R.S.Mishra, "Total Quality Management: model and flexible framework for performance improvement", International Journal of Research and Analytical Reviews (IJRAR), E-ISSN 2348-1269, P- ISSN 2349-5138, Volume.6, Issue 2, Page No pp.867-874, June 2019, DOI: http://doi.one/10.1729/Journal.23462*
50. *Rao, S. Subba, Luis E. Solis, and T. S. Raghunathan. 1999. "A Framework for International Quality Management Research: Development and Validation of a Measurement Instrument." Total Quality Management 10 (7): 1047–1075*
51. *Scannell, T.V., Vickery, S.K., and Droge, C.L., 2000. Upstream supply chain management and competitive performance in the automotive supply industry. Journal of Business Logistics, 21 (1), 23–48.*
52. *Solis, Luis E., T. S. Raghu-Nathan, and S. Subba Rao. 2000. "A Regional Study of Quality Management Infrastructure Practices in USA and Mexico." International Journal of Quality & Reliability Management 17 (6): 597–614*
53. *Slappendel, C. (1996), "Perspectives on innovation in organizations", Organization Studies, Vol. 17 No. 1, pp. 107-29.*
54. *Soltani, E., van der Meer, R., & Williams, T. M. (2005). A Contrast of HRM and TQM Approaches to Performance Management: Some Evidence*. British Journal of Management, 16(3), 211–230. doi:10.1111/j.1467-8551.2005.00452.x*
55. *Seetharaman, A., Sreenivasan, J., & Boon, L. P. (2006). Critical Success Factors of Total Quality Management. Quality & Quantity, 40(5), 675–695. doi:10.1007/s11135-005-1097-2*
56. *Sulphey, M. M. "The utility of Q-methodology in Human Resource Management research." International Journal of Human Resources Management 3.3 (2014): 15-26.*
57. *Saldana Delia (2001) , " cultural competency a practical guide for mental health service providers ", Hogg Foundation for Mental Health , The university of Texas at Austin , pp. 1-30 .*
58. *Saaty,T.L.(1980), The Analytic Hierarchy Process, McGraw-Hill,NewYork,NY.*
59. *Saaty,T.L.(1994), " How to make a decision: The Analytic Hierarchy Process", Interface, Vol.24 No.6,Pp. 19-43.*
60. *Singh, Komal, and AkshayDeokar."Effects of 5S Implementation on Performance of Organization." International Journal of Business and General Management (IJBGM) 7.2 (2018): 1-14.*
61. *Saaty, T. L., & Peniwati, K., (2007).Group decision making. Pittsburgh: RWS.*

62. *Tüysüz, F., & Şimşek, B. (2017). A hesitant fuzzy linguistic term sets-based AHP approach for analyzing the performance evaluation factors: an application to cargo sector. Complex & Intelligent Systems, 3(3), 167–175. doi:10.1007/s40747-017-0044-x*

63. *Vanichchinchai, A., & Igel, B. (2011). The impact of total quality management on supply chain management and firm's supply performance. International Journal of Production Research, 49(11), 3405–3424. doi:10.1080/00207543.2010.492805*

64. *Van Schalkwyk, J. (1998). Total quality management and the performance measurement barrier. The TQM Magazine, 10(2), 124-131.*

65. *Walsh, A., Hughes, H., & Maddox, D. P. (2002). Total quality management continuous improvement: is the philosophy a reality? Journal of European Industrial Training, 26(6), 299–307.*

66. *White, M.A. and Nebeker, D.M. (1996), "Team-oriented performance management: an alternative to traditional performance appraisal", in Knouse, S.B. (Ed.), Human Resources Management Perspectives on TQM, Concepts and Practices, ASQC Quality Press, Milwaukee, WI, pp. 193-215.*

67. *Yildirim, B. F., & Adiguzel Mercangoz, B. (2019). Evaluating the logistics performance of OECD countries by using fuzzy AHP and ARAS-G. Eurasian Economic Review. doi:10.1007/s40822-019-00131-3*

68. *Zbaracki, M.J. (1998) 'The rhetoric and reality of total quality management', Administrative Science Quarterly, pp. 602-636.*

69. *Zadeh, L. (1965). Fuzzy sets. Information Control, 8, 338–353.*

AUTHOR BIOGRAPHIES



Mr. Rakesh Kumar Graduated in Industrial Engineering and Master's in Mechanical Engineering with specialization in Computer Integrated Manufacturing from India. He also obtained PGD in IPR. He is currently an Assistant Professor and research scholar at Delhi Technological University in Department of Mechanical. His research interest includes Total Quality Management, Reliability Engineering, Operations and Supply chain management, Fuzzy Logic, AI and Decision Science and Intellectual Property Rights.



R.S. Mishra (Professor) received the Ph.D. degree in Mechanical engineering from IIT Delhi, India and currently at Delhi Technological University serving since 1997. He having 150 research papers in peer reviewed in International journals and 75 research papers in proceedings of International and National conferences. Prior to joining DTU, he served 09 years at different engineering colleges of India, viz, REC Hamirpur, PAU Ludhiana (Punjab) and HAU, Hisar (Haryana.)

View publication stats

Proceedings of the Second International Conference on Inventive Research in Computing Applications (ICIRCA-2020)
IEEE Xplore Part Number: CFP20N67-ART; ISBN: 978-1-7281-5374-2

Prediction and forecast for COVID-19 Outbreak in India based on Enhanced Epidemiological Models

Siddharth Singh *
Microelectronics Research Laboratory
Department of Applied Physics
Delhi Technological University
New Delhi, India
siddharthsingh16aug@gmail.com

Piyush Raj*
Microelectronics Research Laboratory,
Department of Applied Physics
Delhi Technological University
New Delhi, India
piyushraj.9898@gmail.com

Raman Kumar
Microelectronics Research Laboratory
Department of Applied Physics
Delhi Technological University
New Delhi, India
raman.dtu.a7@gmail.com

Rishu Chaujar
Microelectronics Research Laboratory
Department of Applied Physics
Delhi Technological University
New Delhi, India
chaujar.rishu@dtu.ac.in

Abstract—This paper has analyzed the COVID-19 outbreak in India. This Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV-2) also known Novel Coronavirus or COVID-19 is an infectious and contagious viral Respiratory disease that will analyze the magnitude of outbreak in India and future projections of the same. In China, where the outbreak first started, the basic reproduction number, R_0 is found to be lying between 2-2.5, varying across regions. In India, there have been a total of 70767 confirmed cases of people infected with this disease as reported by the Government of India up to 11 May 2020. In this work, the difference between number of actual reported confirmed cases and approximate number of actual cases, due to insufficient number of tests being conducted, is highlighted based on a unique approximate mathematical formula, thereby establishing relationship between Death Count due to disease and number of people infected with it. Further, utilizing ICMR's available data about COVID-19 patients in India and employing an Enhanced Version of SIR Epidemic Model also known as SIRD devised by generating optimal parameter values and taking number of deaths due to pandemic into account, the time dependence of Outbreak's Intensity in India forecasting maximum number of confirmed active cases of COVID-19 present in a day (Peak Value) and also predicting total number of deaths in India due to the outbreak.

Keywords— COVID-19; SARS-CoV-2; R_o; ICMR; SIRD; Epidemic Model.

* Indicates equal contributors

I. INTRODUCTION

The world is in a crisis situation, currently dealing with a global pandemic in the form of COVID-19. Until 10 May 2020, there have been 274361 deaths registered due to it; globally out of which, 2109 are from India alone as reported by WHO [1]. This work is aimed at revealing certain facts about the current situation not presented clearly by data and also predicting and forecasting about future situations using an enhanced version of SIR Epidemic model explained further in this work. The analysis and modelling has been done in Python Programming language since analysis requires representations in form of multiple graphs.

The global average fatality rate [2][3] which is defined as the percentage of people who have died due to disease, out of the total infected people, has been observed to be around 3.4% as reported by WHO up to March 3 2020. In India, the total confirmed cases of death of people due to COVID-19 until 10 may 2020 is 2213[4]. The fatality rate is less for India as compared to the rest of the world, but at the same time, tests conducted on new individuals in India are less in number as compared to its population. In India, a total of 169037 tests have been conducted as reported by ICMR until 10 May 2020[4].Each individual suspected of being infected by COVID-19 undergoes 2-4 tests [5] and if he is found to be infected, then before recovering and being discharged from the hospital, on an average 6-8 tests [5] are conducted on the person. Although the rate of tests in India has increased, it is still quite low as around 1.41 tests are conducted per thousand people in India [6]. *"Fig. 1,"* shows data of COVID-19 tests conducted on a daily basis on new people and existing patients. It can be figured out that neither reported confirmed

© IEEE 2020. This article is free to access and download, along with rights for full text and data mining, re-use and analysis

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:17:24 UTC from IEEE Xplore. Restrictions apply.

positive cases of infected by COVID-19, nor tests or test positivity rate shown in *"Fig. 2,"* can be used as a completely reliable metric to predict future number of cases and deaths, as there are bound to be asymptomatic patients [7], who recover from disease without being reported or hospitalized and at the same time acting as carriers for the disease [7]. Number of reported deaths due to COVID-19 can give a more legitimate guess of the number of actual cases if a fixed value of fatality rate is assumed and considering the fact that each death due to COVID-19 has been reported. Based on these assumptions, an approximate model has been created

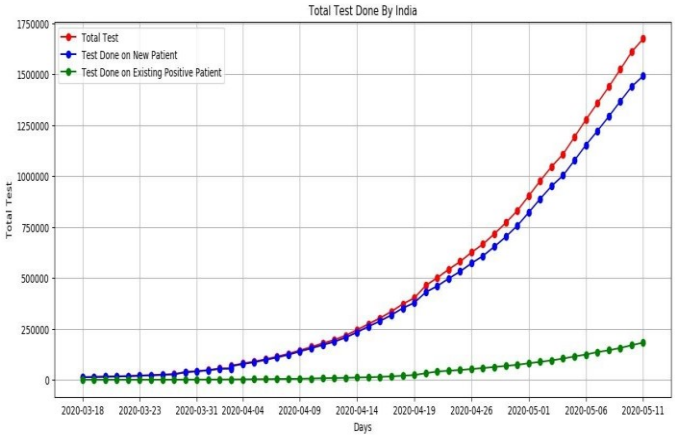


Fig. 1. Daily number of tests conducted in India

Assuming that upon being infected if a person dies and takes on an average y days to decease, to account for slight variations, let it be $y \pm \varrho$. If on n^{th} day, let the total number of deaths be D(n) and T(n) be the total number of actual cases of people freshly infected from COVID-19 ,i.e. new individuals infected from COVID-19 on n^{th} day and further, assuming any constant value of fatality rate as m, then this may be approximated as

$$m\,\%\,of\;T(n-(y\pm\varrho))=D(n) \qquad (1)$$

—— (2)

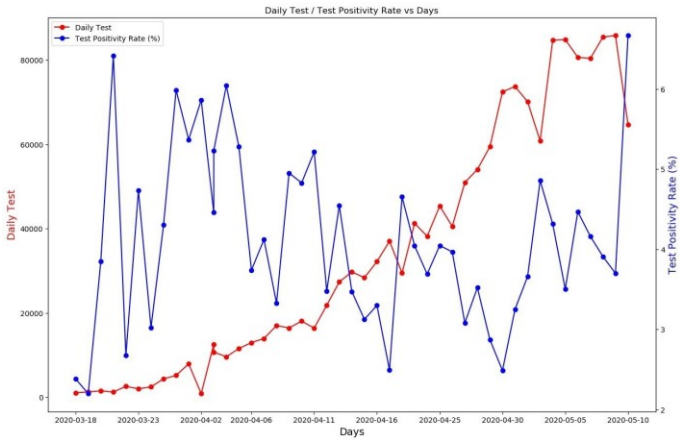


Fig. 2. Plot of daily tests(Red) Conducted and test positivity rate(Blue) for new tests

Here, $T(n-(y\pm\varrho))$ represents number of people infected in $y \pm \varrho$ days before n^{th} day. Here, it is assumed the fresh cases, taken for a day, due to which percent of people died after y days neglecting small variations of ϱ is 'm'. Based on this approximation, the total number of people infected with COVID-19, updated on a daily basis has been plotted in *"Fig. 3,"*for different values of assumed fatality rate m. The incubation period for the disease in which a patient starts to develop symptoms varies between 0 -14 days and the median time taken by a person to die after showing first symptoms is 14 days[2][3] , based on this fact the optimal value of y lies between 16-24.

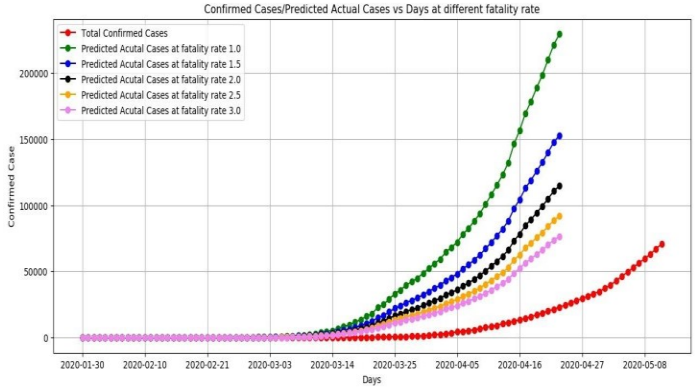


Fig. 3. Plot of approximate predicted cases of COVID-19 vs actual reported cases for assumed different values of fatality rate m and y=18

II. METHODOLOGY

Several epidemic models have been employed to forecast the outbreak of an infectious disease. The most commonly used is SIR (Susceptible-Infectious-Recovered) epidemic model [8] in which the total population (N) is classified into three categories or classes namely people susceptible of being Infected (S), people who are Infected (I) and the people who

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:17:24 UTC from IEEE Xplore. Restrictions apply.

have recovered from disease or have developed immunity against it(R).A closely related another model can be used having another class, which consists of number of deceased persons (D). The model, hence, can be named as Susceptible-Infectious-Recovered-Death (SIRD) model [9]. It is governed by the following set of equations:

(3)

(4)

(5)

(6)

All the four classes are a function of time. It is assumed that Infectious population, I, spread disease to Susceptible population, S. Here, β is transition rate (S →I) which relates susceptible population to infectious population as shown in Equ.3, γ is recovery rate (I →R) which relates recovered population with infected population as shown in "(5)". Out of the infected population, I, D people die due to disease, δ is a mortality rate (I → D) which relates Infectious population to deceased population. Figure 4 depicts the relationship between S,I,R and D in terms of parameters β,γ and δ in SIRD modelling pictorially .Total population N at any given time is assumed to be constant since deaths compared to total population is negligible i.e. N= S+I+R+D and hence

(7)

The basic reproduction number of SIRD model is given by:

(8)

Upon applying the SIRD model, forecasting can be done for future values of infected population I, Susceptible population S, Recovered population R and deceased population D. Based on the equations *"Fig. 4,"* shows relationship between parameters β, δ, and γ with the variable to be determined namely S,I,R and D as a function of time. The forecasting from the model can be represented pictorially with time dependent values of all variables plotted on y axis *"Fig. 5,"* shows a dummy graphical representation of forecasting from the model in terms of fraction of population.

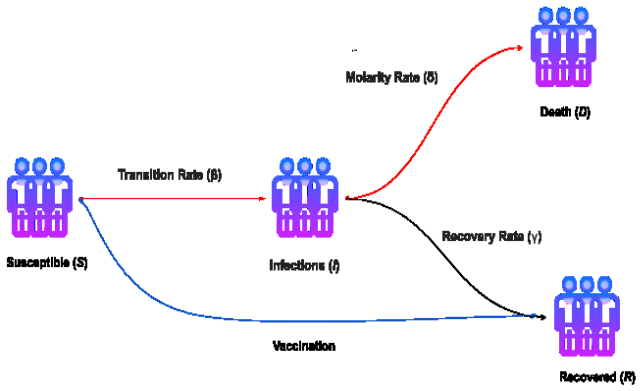


Fig. 4. Pictorial representation of SIRD epidemic model in terms of parameters

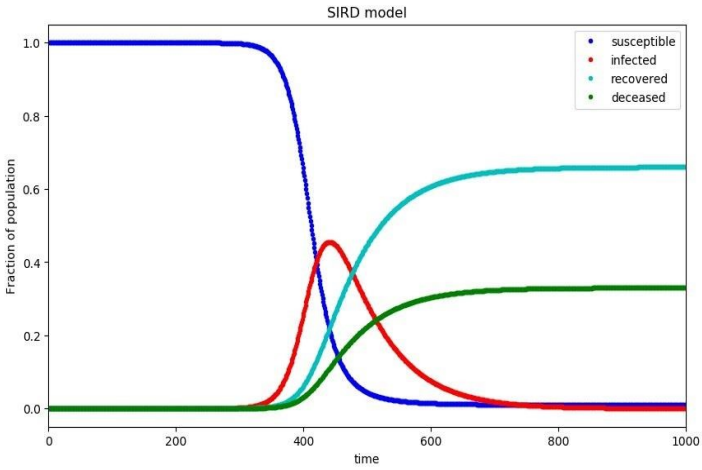


Fig. 5. Dummy graphical representation of forecasting using SIRD epidemic Model

While applying the SIRD epidemic model to predict future cases of people infected and deceased in India based on current available data as provided by ICMR [4], certain considerations were taken. Instead of taking population into account as susceptible class S, only a fraction of population is considered which is based on the fact that the Ministry of Health and Family Welfare, Government of India has divided all districts of the country into three Zones namely Red, Orange and Green. As of May 1 2020 those districts in which not a single case of COVID-19 had been recorded since the past 21 days had been categorized as Green Zone. A total of 319 districts had been put in this category out of the total 720 districts [4], approximately around 300 districts are such where not a single case of person being infected with COVID-19 was ever recorded nor death due to the same. Outbreaks have occurred in rest of zones, severely in red Zone districts and mild in Orange Zone districts. Districts in India are

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:17:24 UTC from IEEE Xplore. Restrictions apply.

divided based on population density and not on land area, so it is safe to assume that approximately 40% of population out of total population of "1.32 Billion"[10] can be neglected while assuming value of N. *"Fig. 6,"* shows zone wise distribution of districts in India.

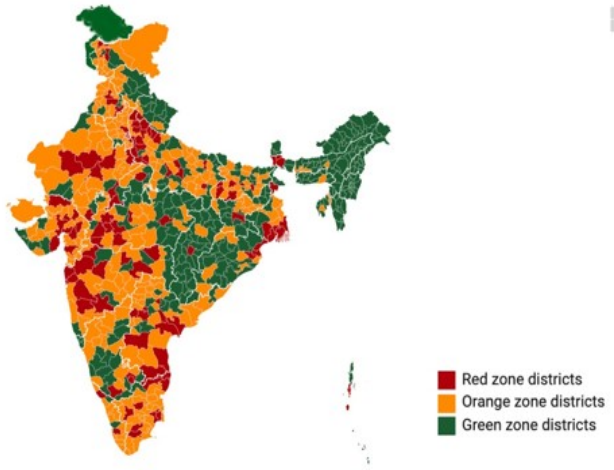


Fig. 6. Map of India showing districts categorized into different Zones [11]

Three different variants of the SIRD model were created to get the best values of forecast.

- In the first variant, the mortality rate δ was fixed and values of parameter β and γ were determined through curve fitting on available data of infected patients.
- In the second variant the normal SIRD model was applied, values of parameters δ, β and γ were determined through curve fitting on available data of infection.
- The mortality rate δ was linked to the recovery rate γ in the third variant and value of parameters β and γ were determined through curve fitting on available data of infected patients. Once γ was determined the mortality rate was calculated.

After trying these three different variants of the SIRD model, it was observed that combination of the 2nd and 3rd variant gave the best results. The model was fitted with available data of infection up to 11 May 2020 to determine the value of parameters δ, β and γ. Once values were determined, a graph was plotted to forecast the future cases of infection among people from COVID-19. Initial values of S,I,R,D plays an important role in prediction using this model.

III. RESULTS AND DISCUSSION

The predicted and actual reported cases of COVID-19 in India based on data up to 11 May 2020 coincides up to a certain extent as has been shown in *"Fig. 7,"* . Slight variations can be accounted for by the fact that testing and test positivity was not uniform and that reported cases were not equal to actual cases. Hence, the model can yield reliable results for further forecasting. *"Fig. 8,"* shows forecasting results of SIRD model for next 100 days which reveals certain facts:

- The total Expected maximum number of active cases of people infected from COVID-19 in a day(Peak Value) will be 825058
- The Peak value of outbreak is expected to occur on 12 June 2020
- The predicted value of basic reproduction number R_o from the model is 1.12
- The total number expected deaths of people infected due to the disease is 36941

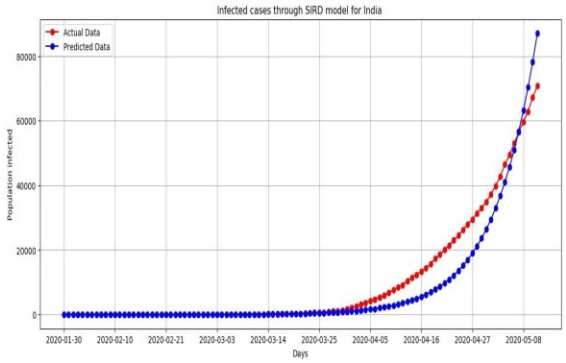


Fig. 7. Graph depicting predicted and actual reported cases of COVID-19 in India upto 11 May 2020

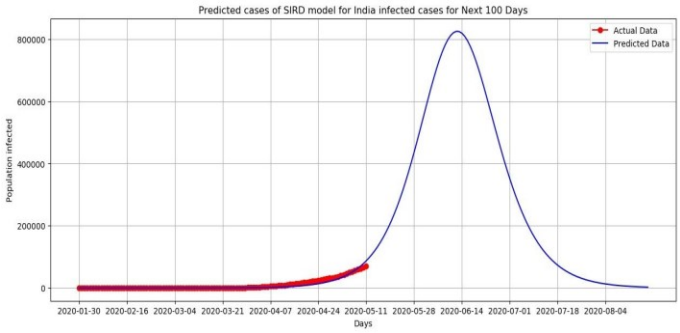


Fig. 8. Graph depicting predicted cases of COVID-19 in India using SIRD Model

Based on the modelling, it can be observed from figure 9 that the number of susceptible people being infected from COVID will decrease and the situation will improve by the month of August and September. It can be seen that cases of infected people will rise gradually at comparatively lower slope but slope will keep on increasing and cases will rise exponentially up to an extent till it reaches a max value (Peak Value), then number of infected people will start to decrease

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:17:24 UTC from IEEE Xplore. Restrictions apply.

gradually and finally in the mid-august or September outbreak will be contained.

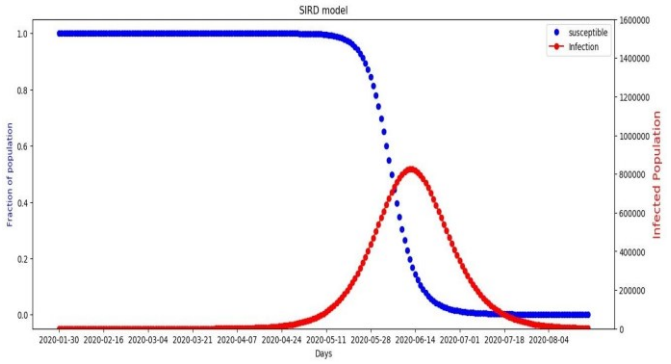


Fig. 9. Graph depicting variation in susceptible population and infected population in India based on SIRD Model

This modelling is based on the current data available, and empathetically it is believed that there is bound to be a certain extent of unreliability. Also under most of this period, complete India was in lock down condition with multiple restrictions especially on travel, which are bound to be released and hence, conditions will also change and chances of spread of this highly contagious disease will increase and value of R_o will also increase drastically since social distancing norms would not be followed everywhere. No vaccine or cure to this disease is still available, in the near future if a cure is found or a vaccine is developed, the situation may change drastically. Hence results mentioned here should only be used to understand the situation quantitatively based on the condition of lock down and the restrictions imposed by it.

IV. CONCLUSION

This work has tried to forecast the situation of India amid the outbreak of global pandemic COVID-19.The outbreak of COVID-19 in India was analyzed using approximate mathematical modelling and then by an enhanced version of the SIR epidemic model. Instead of assuming values of parameters for SIRD modelling, sophisticated methods were applied implementing curve fitting on existing data and linking mortality rate δ to the recovery rate γ, in order to obtain more realistic results. The initial value of susceptible patients was based on Zone bifurcation implemented by the Government of India to get optimal results. It has been observed that testing of susceptible people in India is insufficient as compared to population. Low value of R_o and lesser value of fatality rate suggests that spread of disease in India is less fatal as compared to other parts of the world as has been revealed by the model. Model predicts that over 35000 deaths may occur due to the outbreak and a maximum of around 800000 active cases may be present on a single day. Although the current situation based on available data seems to be acceptable as compared to other parts of world. This has been made possible probably due to timely enactment of lock down and strict implementation of social distancing norms. But due to relatively very high population even at lower values of R_0 , number of infected patients may be very high and even at lower fatality rate huge numbers of deaths may occur due to COVID-19 in India has been predicted by model. Containment of disease based on the model is based on the assumption that conditions will remain the same, and cases predicted may not match actual confirmed cases due to inadequate number of tests and presence of a large number of asymptomatic patients.

REFERENCES

[1] World Health Organisation Novel CoronaVirus 2019 (COVID-19)

[2] Battegay Manuel, Kuehl Richard, Tschudin-Sutter Sarah̲a̲, Hirsch Hans H, Widmer Andreas F, Neher Richard A . 2019-Novel Coronavirus (2019-nCoV): estimating the case fatality rate – a word of caution .Swiss Med Wkly. 2020;150:w20203 doi : 10.4414/smw.2020.20203

[3] A. C. Ghani, C. A. Donnelly, D. R. Cox, J. T. Griffin, C. Fraser, T. H. Lam, L. M. Ho, W. S. Chan, R. M. Anderson, A. J. Hedley, G. M. Leung .Methods for Estimating the Case Fatality Ratio for a Novel, Emerging Infectious Disease. American Journal of Epidemiology, Volume 162, Issue 5, 1 September 2005, Pages 479–486, doi: 10.1093/aje/kwi230

[4] Ministry of Health and Family Welfare Government of India

[5] National Centre for Disease Control is an institute under the Indian Directorate General of Health Services, Ministry of Health and Family Welfare.

[6] Joe Hasell, Esteban Ortiz-Ospina, Edouard Mathieu, Hannah Ritchie, Diana Beltekian, Bobbie Macdonald and Max Roser. Statistics and Research Coronavirus (COVID-19) Testing

[7] Yan Bai, MD[1]; Lingsheng Yao, MD[2]; Tao Wei, MD[3]; Fei Tian, MD[4]; Dong-Yan Jin, PhD[5]; Lijuan Chen, PhD[1]; Meiyun Wang, MD, PhD[1]. Presumed Asymptomatic Carrier Transmission of COVID-19.JAMA.2020;323(14):1406-1407.doi:10.1001/jama.2020.2565

[8] Binti Hamzah, Fairoza Amira & Hau, Cher & Nazri, Hafeez & Ligot, Dominic & Lee, Guanhua & Shaib, Mohammad & Zaidon, Ummi & Abdullah, Adina & Chung, Ming & Ong, Chin & Chew, Pei. (2020). CoronaTracker: World-wide COVID-19 Outbreak Data Analysis and Prediction. [Preprint]. Bull World Health Organ. E-pub: 19 March 2020. doi : 10.2471/BLT.20.255695.

[9] Cleo Anastassopoulou, Lucia Russo, Athanasios Tsakris, Constantinos Siettos. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. Plosone March 31, 2020 doi: 10.1371/journal.pone.0230405

[10] Ministry of Statistics and Programme Implementation Government of India.

[11] Health Analytics Asia Mapping Covid 19 Hotspots In India

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:17:24 UTC from IEEE Xplore. Restrictions apply.

Soft Computing
https://doi.org/10.1007/s00500-020-05295-8

METHODOLOGIES AND APPLICATION

QNBP NN-based *I cos* ϕ algorithm for PV systems integrated with LV/MV grid

Priyanka Chaudhary[1] · M. Rizwan[1]

© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract
In this paper, a robust control technique is designed presented for a grid-integrated solar photovoltaic system. The proposed system consists of a solar photovoltaic array, converter along with linear/nonlinear loads. The main aim of this controller is to obtain maximum power from solar photovoltaic array and control the DC link voltage and injected grid current under different scenarios. The performance analysis of the proposed controller is demonstrated under different scenarios and meteorological conditions. The proposed system performs very well under transient conditions and gives very fast response. Moreover, various power quality indices were monitored and analyzed at unity power factor using hybrid control technique based on *I cos* ϕ approach with quasi-newton backpropagation learning. The grid current distortion is kept within the prescribed limits given by IEEE-519 and the IEEE-1547 standard. Experimental results illustrate the capabilities of proposed system under abnormal and unbalanced conditions.

Keywords Solar PV systems · Inverter · Maximum power point tracking (MPPT) · Synchronization · Power quality

1 Introduction

The solar photovoltaic technology is widely utilized for distributed power generating systems. Further, it has a potential to fulfill the demand when exploited on a large scale. However, there are certain issues and challenges associated with the integration of such highly efficient and potential resource with the grid (Karimi et al. 2016; Chaudhary and Rizwan 2018a). Efficient controllers are required in order to achieve maximum power output since the meteorological parameters highly affect the PV output. The controller must be able to provide fast response under transient conditions also. Increase in the demand of solar PV systems at rural level is due to ease in installation, no cost of fuel, and low maintenance cost. Also, the solar PV systems are gaining popularity in urban areas as well due to numerous advantages like peak load reduction and reduction in electricity charges. Various organizations government and nongovernment associated with solar PV business are continuously supporting the installation solar PV system at rooftop for uninterrupted power supply. Moreover, the power generated from solar photovoltaic-based power plants is varying due to sudden change in meteorological parameters, and therefore, it is mandatory to provide a proficient controller to obtain the maximum power from plants under varying meteorological and atmospheric parameters, and as a result, the integration of solar photovoltaic based power plants would be done efficiently and negative impacts like large fluctuations of voltage at a point of common coupling and grid stability may be reduced (IEEE Std. 519 1992). Aforesaid standards are responsible for maintaining the grid stability and power quality. Therefore, large-scale deployment of solar PV-based generation technologies increases the need of accurate and coordinated controllers. These controllers should be capable enough for efficient accommodation of solar PV power generating systems into the grid with maintaining the power quality standards.

Solar PV systems are being availed on a massive scale to fulfill the demand as well as helping the grid as well to mitigate the power quality problems. Single-stage and two-stage systems are being used to integrate the solar PV systems with the grid as reported in the literature (IEEE

Communicated by V. Loia.

✉ Priyanka Chaudhary
priyankach.iilm@gmail.com

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

Published online: 25 September 2020

Std. 519 1992). Since this is a very emerging area of research in the area of solar photovoltaic system, a lot of research is being carried out by esteemed researchers and developed and presented in the literature (Chaudhary and Rizwan 2015, 2017, 2018b; Rajput et al. 2017; Kumar and Mishra 2014; Arya et al. 2015; Singh et al. 2014; Chen et al. 2009; Mahmud et al. 2014; Agarwal et al. 2017; He and Li 2013; Singh et al. 2016; Reisi et al. 2013; Masoum et al. 2002; Rizwan et al. 2014). Recently, lot of work have been going on in the area of adaptive controllers for solar photovoltaic systems in addition to many well-established conventional techniques. Further, a controller to provide an adjustable DC link voltage at PCC using grid supported solar photovoltaic plant was proposed (Chaudhary and Rizwan 2018b). The proposed system could incorporate the sudden changes at point of common coupling is developed and presented. It is also shown by various researchers that for a three-phase system, single stage configuration is more advantageous. However, conventional single-stage SPV systems suffer from the drawback of converters being idle during unavailability of solar PV power. An optimized control algorithm JAYA to optimize the coefficients of PI controller and filter parameters of PV fed DSTATCOM has been proposed (Rajput et al. 2017). Researchers are involved with the development of robust controllers for performance enhancement of three-phase grid-connected solar PV systems (Kumar and Mishra 2014; Arya et al. 2015; Chaudhary and Rizwan 2017; Singh et al. 2014). The researcher introduced a new control technique based multi-functional distribution static compensator that works in voltage-control mode in rigid source. The voltage regulations provided by this method at the time of voltage disturbance on load terminals, is fast. Additionally, at the time of normal operations, source current's control is allowed by the developed reference load voltage. Therefore, for making source power factor to the unity, distribution static compensator adds up the re-active and harmonic component of load current (Kumar and Mishra 2014). In this paper, a biogas/biomass diesel engine which operates as a prime mover to distributed power generator, is driving and controlling the synchronous reluctance generator (SyRG), given generator operates as source to give linear and nonlinear loads. A DSTATCOM that depends on a three-leg voltage source converter is carried out adopting a control technique based on the neural network, for suppressing harmonics, balancing loads, and regulating voltage in three-phase SyRG with battery energy storage system (BESS) (Arya et al. 2015). This paper presents a control algorithm that depends on the predictive current control for voltage-source-inverter fed with a solar PV system, which is further integrating to distribution grids with perturb and observe maximum power point tracking (MPPT) algorithm. Given control technique is used on a 3-phase system for supporting the utility grid with giving power to the grid along with the loads that are connected with enhanced power quality. This used control method on solar PV voltage-source-inverter gives the quick transitory responses and enhanced control accuracy (Chaudhary and Rizwan 2017). This paper investigates the various DSTATCOMs topologies, the art's state and performances, consideration of designing, future improvements, and capable applications for improving the quality of power. For various operations, i.e., compensating the re-active power, eliminating harmonic, balancing of loads, and compensating the neutral current, development and installation of given distribution static compensators for three-phase three-wire system and three-phase four-wire systems takes place. The goal of this paper is to deeply prospecting the DSTATCOMs to various researchers, engineers, and the communities that deals with power quality enhancement. The researcher has also provided the classified list of current research publications (Singh et al. 2014). A robust fuzzy-controller for solar PV fed inverter is proposed in Chen et al. (2009) for the stable operation of grid-connected solar PV system and robustness of controller is achieved by using the Taguchi tuning algorithm. A new nonlinear current control scheme for grid-connected photovoltaic (PV) system using partial feedback linearization is presented in Mahmud et al. (2014). The proposed controller linearizes the system partially and calculate reference current maximum power point tracking (MPPT) system. In this paper, three-phase single-stage grid-tied solar photovoltaic energy conversion system (SPECS) with phase-locked loop-less (PLL-less) fast character of triangle function (CTF) control algorithm is presented. The used setup constructs a SPV vector, voltage-source-inverter, ripple-filter and three-phase grids using linear nonliner loads connection. Solar photovoltaic energy conversion system not only supplies active power to grid and linked loads, it also gives performances of an active shunts filter, i.e., compensating re-active power, attenuating harmonic, power factor correction, balancing of loads and ZVR (Agarwal et al. 2017). The given procedure grants the coordination of closed-loop control of DG fundamental voltage and line harmonic current. By using HCM, loads of local harmonic can be satisfied without extracting harmonic current. Additionally, HCM assures the smooth conversion at the time of grid-connected/islanding operating mode transmission (He and Li 2013). A two-stage, three-phase grid interface SPV energy conversion along with an adaptive notch filter-depended control technique is presented through this paper. The considered SPV is multi-functional grid interface solar photovoltaic energy conversion system, as well as converting DC power to AC from SPV, and has an ability of compensating re-active power, eliminating harmonic current and balancing the

loads in three-phase active current distribution system (Singh et al. 2016).In this paper, the single-stage, three-phase grid-connected SPV is being presented. The given system is capable enough for both, feeding the extracted solar energy to grid as well enhancing the quality of power in distribution system. Purposes of this given system are MPPT, feeding of SPV energy into grid, and mitigating harmonic of load linked at point of common coupling (PCC) and grid current balancing. A three-phase voltage-source-converter is used by the solar photovoltaic system to perform all the operations. For controlling the VSC, an improved linear sinusoidal tracer (ILST)-based control technique is used. In this, a varying direct current link voltage has been applied for MPPT. For quick dynamic feedback, instantaneous compensatory technique is being used that incorporates the change in photovoltaic power (Sing et al. 2014). Researchers in Kannan and Rengarajan (2014) developed a control algorithm based on *I cos* ϕ for DSTATCOM. The proposed controller provides harmonic elimination in source current effectively as well as provide reactive power compensation. The authors used a fuzzy logic controller (FLC) for DC link voltage control instead of conventional proportional integral (PI) controller. Control approaches using artificial neural networks (ANN) are reported by researchers for numerous applications (Kumar and Mangaraj 2013; Jain and Singh 2016; Arya et al. 2014; Mazumdar and Harley 2008; Arya and Singh 2014; Singh and Arya 2014; Singh et al. 2011; Jayachandran and Sachithanandam 2015; Chaudhary and Rizwan 2018c; Mangaraj and Panda 2017; Janpong et al. 2011; Kulkarni and Venayagamoorthy 2009; Varma et al. 2015) but less available for solar PV integration with grid related application. A research given in Hamid et al. (2014) poses a study to lower the current harmonics from a photovoltaic (PV) plant. Power conditioner unit is used by this system that drives in feed-forward mode and placed parallelly to the plant. For the sinusoidal flow of total current in the grid, PCU compensates the output current exaggeration of the PV plant. An efficient control structure is used by system to fuses a sine-wave formation for and components extraction of distortion current and conditioner current control. To execute the functions, i.e., harmonic mitigation, power factor correction under reactive loads that lowers the DC link voltage in capacitor of voltage source converter (VSC), a gradient descent back propagation for three-phase two-level distribution static compensator is presented in Kumar and Mangaraj (2013). To satisfy the linear/nonlinear loads researchers presented a system in Arya et al. (2014) this system uses the DSTATCOM by practicing the anti-Hebbian control technique. Control techniques are being used by the researcher for deriving fundamental active and reactive power components of load current to conclude supply current. This technique is enforced on a DSTATCOM applying digital signal processor (DSP) for reactive power compensation, harmonics elimination, and load balancing.

There are various complications challenging the power utilities in power system, i.e., harmonic's formation and presence of waveform deterioration. If applied voltage waveform on the load is disfigured by another load and not purely sinusoidal, then the obtaining of harmonic current that is inserted to a power network from a nonlinear load will be difficult. The researcher gives the solution to this problem through this paper. To determine the way to differentiate between load harmonics and supply harmonics without detaching the load from the network, neural network is trained for back propagation with the help of time training algorithm (Mazumdar and Harley 2008). Estimation of effecting factors that are responsible for the cause of reduction of power quality issues in distribution system is necessary shown in Arya and Singh (2014). The goals of the trained neural network in DSTATCOM are enhancing the efficiency and keep a track on the capabilities as per the requirement. The research implements the three-phase DSTATCOM applying backpropagation control technique for the tasks, i.e., eliminating harmonic, load stabilizing and compensating reactive power for PFC and ZVR in nonlinear loads in Singh and Arya (2014). Backpropagation depended control technique applied for extracting the values of active–reactive power elements of load current that is needed to estimate reference source current. The blueprint of distribution static compensator is made applying DSP, hence study the completion under several operating conditions (Singh and Arya 2014). To mitigate the power quality problem, DSTATCOM is the optimistic choice of power quality compensator in three-phase four-wire distribution system. PQ compensator's performances depends on control strategies under changing loads and non-ideal sources conditions. The neural dependent power quality control technique is being used by the researcher for distribution static compensator that is made up of four-leg voltage source converter with direct-current capacitor. For converting non-ideal voltage source to ideal sinusoid voltage, extracting the direct current factors of loads, maintaining voltage along capacitor and mitigating neutral current, the used control technique applies five artificial neural network controllers (Jayachandran and Sachithanandam 2015). The main purpose of this research is to analyze the applications of neural network in shunt-active-power-filter. The researchers presented, three out of four components of shunt active power filter, i.e., detecting harmonic, compensation of current, and direct current bus

voltage, has selected a few of neural network architectures as a part of its components (Janpong et al. 2011).

In the proposed research, a proficient controller is designed which can incorporate the various issues such as compensation of reactive power and harmonics related to grid integration of solar photovoltaic based power plants. To obtain and extract the maximum power from the plant, an algorithm namely incremental conductance is utilized here. This algorithm would incorporate the changes due to variations in the input solar energy. Moreover, the simulation results are validated from laboratory-based experimental setup.

2 System description

The proposed solar PV power generation system comprises of PV array, voltage source converter (VSC), MPPT controller, ripple filters, various loads (linear/nonlinear) and AC distribution grid. Figure 1 shows the schematic of the proposed system. Detailed mathematical modeling of the proposed system is presented in this section. As a PV cell is the basic unit of the solar PV array, detailed modeling is shown with the help of equations. A rating of 10 kW_p is chosen for the developed system and a Vikram solar ELDORA 270 of 270.66 W_p for designing the system in MATLAB. Moreover, a solar PV simulator is used to emulate the practical PV characteristics for hardware prototype.

The proposed system is consisting of the following main components.

- Solar photovoltaic array of 10 kW_P (Make: Vikram solar)
- Voltage source converter of 10 kVA rating with operating voltage of 415 V (Make: Semikron)
- Ripple filter having R_f and C_f of 5 Ω, 10 μF respectively.
- Different types of loads and along with utility grid.

The output current of a solar PV cell is given as Pandiarajan et al. (2011):

$$I_{PV} = I_{sc}N_p - N_sI_o\left[\exp\left\{\frac{q(V_{PV} + I_{PV}R_s)}{N_sAkT}\right\} - 1\right]V_{PV} + \frac{I_{PV}R_s}{R_p} \quad (1)$$

Reverse saturation current I_{rs} is given as:

$$I_{rs} = I_{scref} + \left[\exp\left(\frac{qVoc}{NskAT}\right) - 1\right] \quad (2)$$

Saturation current (I_o) as a function of temperature given as:

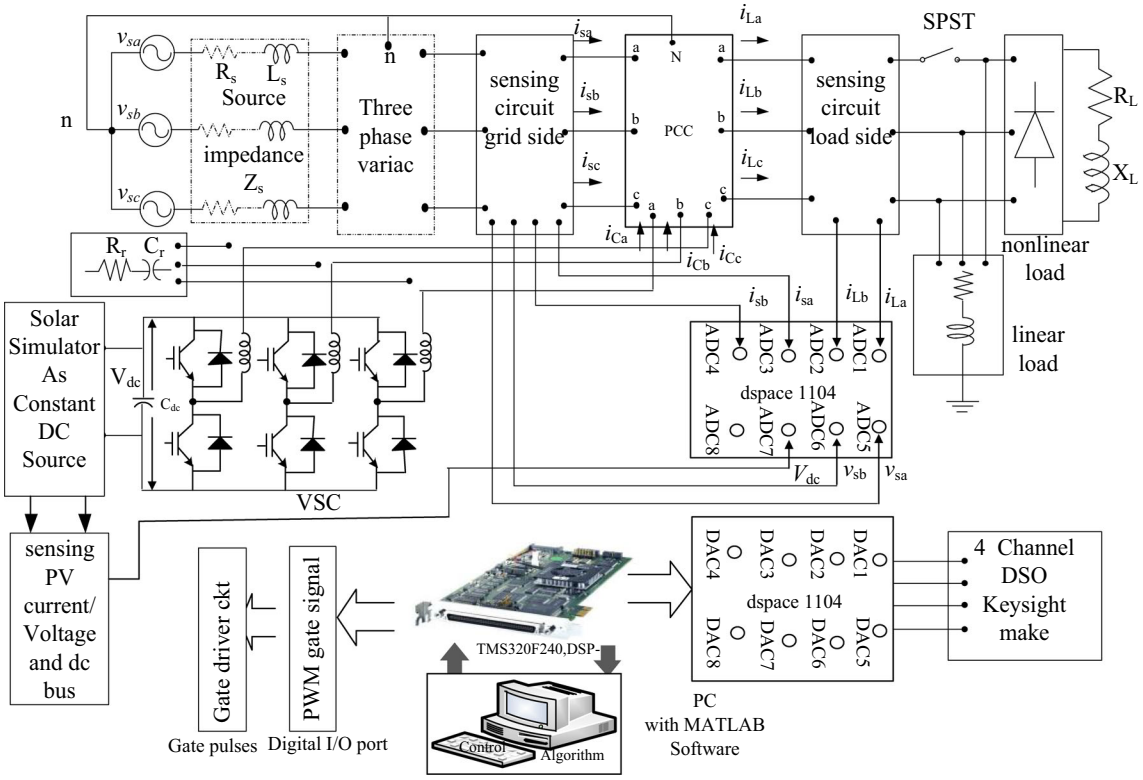


Fig. 1 Schematic of system

$$I_o = I_{rs}\left[\frac{\left(\frac{T}{T_{ref}}\right)^3 e^{qCg}}{Ak} * \left(\frac{1}{T_{ref}} - \frac{1}{T}\right)\right] \tag{3}$$

3 Controller design

In the given solar photovoltaic system, an inverter is used to inject the active power to the grid for power balancing purpose. The solar photovoltaic voltage and frequency are synchronized with the grid voltage and frequency using active synchronization control approach. Following sub-sections depicts the detail modeling of the controller.

3.1 MPPT control

Here, the main aim of the study is to extract maximum possible power from solar photovoltaic array by means of MPPT controllers under varying parameters. A large number of MPPT techniques are reported. Here, incremental conductance approach is utilized to obtain the maximum power by adjusting the reference voltage or reference current signal of the solar photovoltaic array. In this approach, the maximum power point (MPP) is obtained when the slope of the P–V characteristics becomes zero. In case of positive values of slope, the power generated from the solar photovoltaic system is less than the MPP, whereas it is more in case of negative values of the slope. The I–V and P–V characteristics of the solar photovoltaic system at various solar insolation is provided in Fig. 2.

$$\frac{dP}{dV} = I + v\frac{dI}{dV} = 0 \tag{4}$$

$$\frac{dI}{dV} \cong \frac{\Delta I}{\Delta V} = -\frac{I_{PV}}{V_{PV}} \tag{5}$$

$$\frac{dP}{dV} = 0 \quad \frac{\Delta I}{\Delta V} = -\frac{I_{PV}}{V_{PV}} \text{ at MPP}$$

$$\frac{dP}{dV} > 0 \quad \frac{\Delta I}{\Delta V} > -\frac{I_{PV}}{V_{PV}} \quad \text{left side of MPP}$$

$$\frac{dP}{dV} < 0 \quad \frac{\Delta I}{\Delta V} < -\frac{I_{PV}}{V_{PV}} \quad \text{right side of MPP}$$

At maximum power point (MPP), output power from solar PV array is the maximum power supplied from the solar PV array.

$$P_{PV} = P_{ref} = V_{PV} I_{PV} \tag{6}$$

$$P_{ref} = P = \frac{3}{2} E_q I_{qref} \tag{7}$$

From Eqs. (28) to (29)

$$I_{qref} = \frac{2}{3}\frac{P_{ref}}{E_q} \tag{8}$$

$$V_{PVref} = \frac{P_{ref}}{I_{PV}}; \quad V_{PVref} = V_{dcref} \tag{9}$$

These reference values will be the input to the inverter controller.

3.2 PV inverter controller

The proposed QNBP-based control algorithm is used to estimate the reference grid currents. The proposed back-propagation technique trained the weights in three-stages, i.e., input signal training's forward feeding, calculating and backpropagating error signals, and then upgradation of training weights. The used algorithm may have one layer or even more. Basically, the proposed algorithm is characterized by three features, i.e., continuation, differentiation,

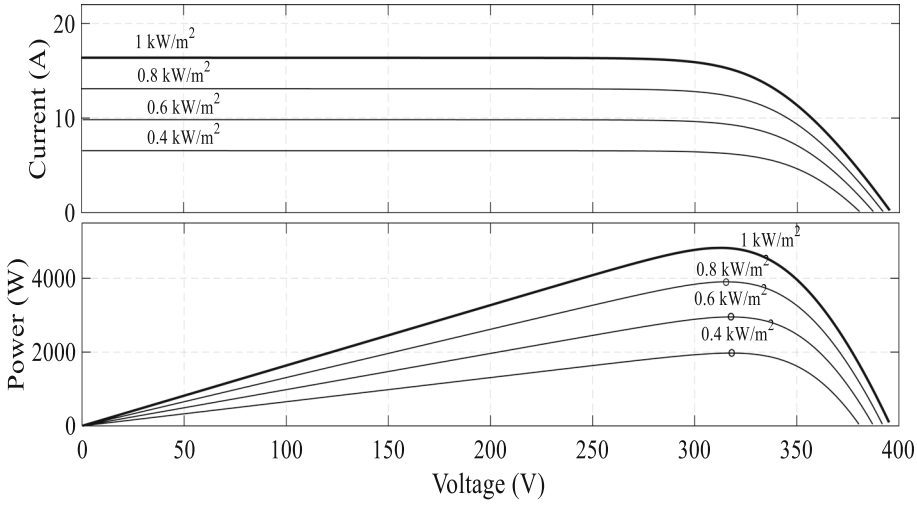


Fig. 2 P–V and I–V characteristics of SPV system at different solar irradiances

and non-decreasing monotony. This algorithm depends on formulas of mathematics and doesn't require any special feature for the functioning of learning practice. Because of mass updation feature on weights, this algorithm provides smooth alterations on the weight's corrections. At the time of training process, the algorithm gets slower because of large amount of learning steps, but after weight's training, the given technique generates notably quick trained output responses. Neural network techniques have more advantage among several control techniques, i.e., feed-forwarding, supervising, reinforcement training and self-organizing principle. Hence, suitable technique for analyzing the linear and nonlinear nature of loads within distribution systems, neural network with perceptron's multilayers, is neural network. The used algorithm is termed as QNBP neural network that include several layers, i.e., input's layer, field layer, and output's layer. The field layer is middle layer which also known as hidden layer. The practical load currents consists of several harmonics, therefore. Hence, the main purpose of inverter controller is to separate the fundamentals load current components. The controller should possesses the capability to filter out DC-offset, fast convergence and small oscillations under steady-state and dynamic conditions. Further, the developed controller training and modifications of neural network weights uses QNBP method (Chaudhary and Rizwan 2019). The proposed controller utilizes to compute weighted value of fundamental active load current component (w_{ap}, w_{bp} and w_{cp}) and fundamental reactive load current component (w_{aq}, w_{bq} and w_{cq}). Moreover, the computed elements produces the reference signals in terms of active and reactive component for source side current.

3.2.1 Calculation of active and reactive loss components

Active loss component is obtained by sensing and comparing DC link voltage (V_{dc}) to reference dc link voltage (V_{dc}^{*}). A PI based controller helps in reducing the error (E_{dc}) between V_{dc} and V_{dc}^{*}. This also helps in regulating the dc link voltage and the outcome of PI controller can be taken as active current component $\left(w_{\mathrm{pdc}}\right)$ (Fig. 3).

$$E_{\mathrm{dc}}(k) = V_{\mathrm{dc}}^{*}(k) - V_{\mathrm{dc}}(k) \tag{10}$$

The controller's output at kth instant is computed as:

$$w_{\mathrm{pdc}}(k) = w_{\mathrm{pdc}}(k-1) + k_{\mathrm{pd}}\{E_{\mathrm{dc}}(k) - E_{\mathrm{dc}}(k-1)\} + k_{\mathrm{id}}E_{\mathrm{dc}}(k) \tag{11}$$

Further, the AC terminal voltage V_{t} and reference terminal voltage V_{t}^{*} is compared and the error between them, processed with the help of a PI controller. Next, this error is further used to obtain reactive loss component. This calculated component further used to maintain the value of AC terminal voltage constant.

$$E_{\mathrm{te}}(k) = V_{\mathrm{t}}^{*}(k) - V_{\mathrm{t}}(k) \tag{12}$$

The controller output at kth instant is given as:

$$w_{\mathrm{qt}}(k) = w_{\mathrm{qt}}(k-1) + k_{\mathrm{pt}}\{V_{\mathrm{te}}(k) - V_{\mathrm{te}}(k-1)\} + k_{\mathrm{it}}V_{\mathrm{te}}(k) \tag{13}$$

A function relating solar PV power (P_{PV}) and AC terminal voltage (V_{t}) has also been implemented in order to obtain fast dynamic response of developed control approach.

$$w_{\mathrm{PV}}(k) = \frac{2P_{\mathrm{PV}}(k)}{3V_{\mathrm{t}}} \tag{14}$$

3.2.2 Estimation of fundamental components

Input layer of neural network (NN) is formed with the sensed values of given load currents (i_{La}, i_{Lb}, i_{Lc}). Here, w_{i} is assumed as initial weight for further calculating to evaluate the fundamental elements of active and reactive load current component. Hence the values of weighted components for all phases are produced with the help of basic $I\ cos\ \phi$ technique:

$$\left.\begin{aligned} i_{Lap} &= i_{La}\mathrm{Cos}\phi_{la}u_{pa} + i_{Lb}\mathrm{Cos}\phi_{lb}u_{pb} + i_{Lc}\mathrm{Cos}\phi_{lc}u_{pc} + w_i \\ i_{Lbp} &= i_{La}\mathrm{Cos}\phi_{la}u_{pa} + i_{Lb}\mathrm{Cos}\phi_{lb}u_{pb} + i_{Lc}\mathrm{Cos}\phi_{lc}u_{pc} + w_i \\ i_{Lcp} &= i_{La}\mathrm{Cos}\phi_{la}u_{pa} + i_{Lb}\mathrm{Cos}\phi_{lb}u_{pb} + i_{Lc}\mathrm{Cos}\phi_{lc}u_{pc} + w_i \end{aligned}\right\} \tag{15}$$

In the above equation $u_{pa,}\ u_{pb}, u_{pc}$ are in-phase unit templates of phase voltagesv_a, v_b and v_c. Line voltages (v_{ab}, v_{bc}, v_{ca}) at PCC are sensed to calculate phase voltages (v_a, v_b and v_c).

$$v_a = \frac{2v_{ab} + v_{bc}}{3},\quad v_b = \frac{-v_{ab} + v_{bc}}{3},\quad v_c = \frac{-v_{ab} - 2v_{bc}}{3} \tag{16}$$

The PCC terminal voltage amplitude (V_{t}) given as:

$$V_{\mathrm{t}} = \sqrt{\left[\frac{2\left(v_a^2 + v_b^2 + v_c^2\right)}{3}\right]} \tag{17}$$

The in-phase unit templates are given as:

$$u_{pa} = \frac{v_a}{V_{\mathrm{t}}},\quad u_{pb} = \frac{v_b}{V_{\mathrm{t}}},\quad u_{pc} = \frac{v_c}{V_{\mathrm{t}}} \tag{18}$$

Further in-phase templates are used to quadrature unit templates.

$$u_{qa} = \frac{u_{pc}}{\sqrt{3}} - \frac{u_{pb}}{\sqrt{3}},\quad u_{ab} = \frac{\sqrt{3}u_{pa}}{2} + \frac{u_{pb} - u_{pc}}{2\sqrt{3}},\quad u_{qc} = \frac{u_{pb} - u_{pc}}{2\sqrt{3}} - \frac{\sqrt{3}u_{pa}}{2} \tag{19}$$

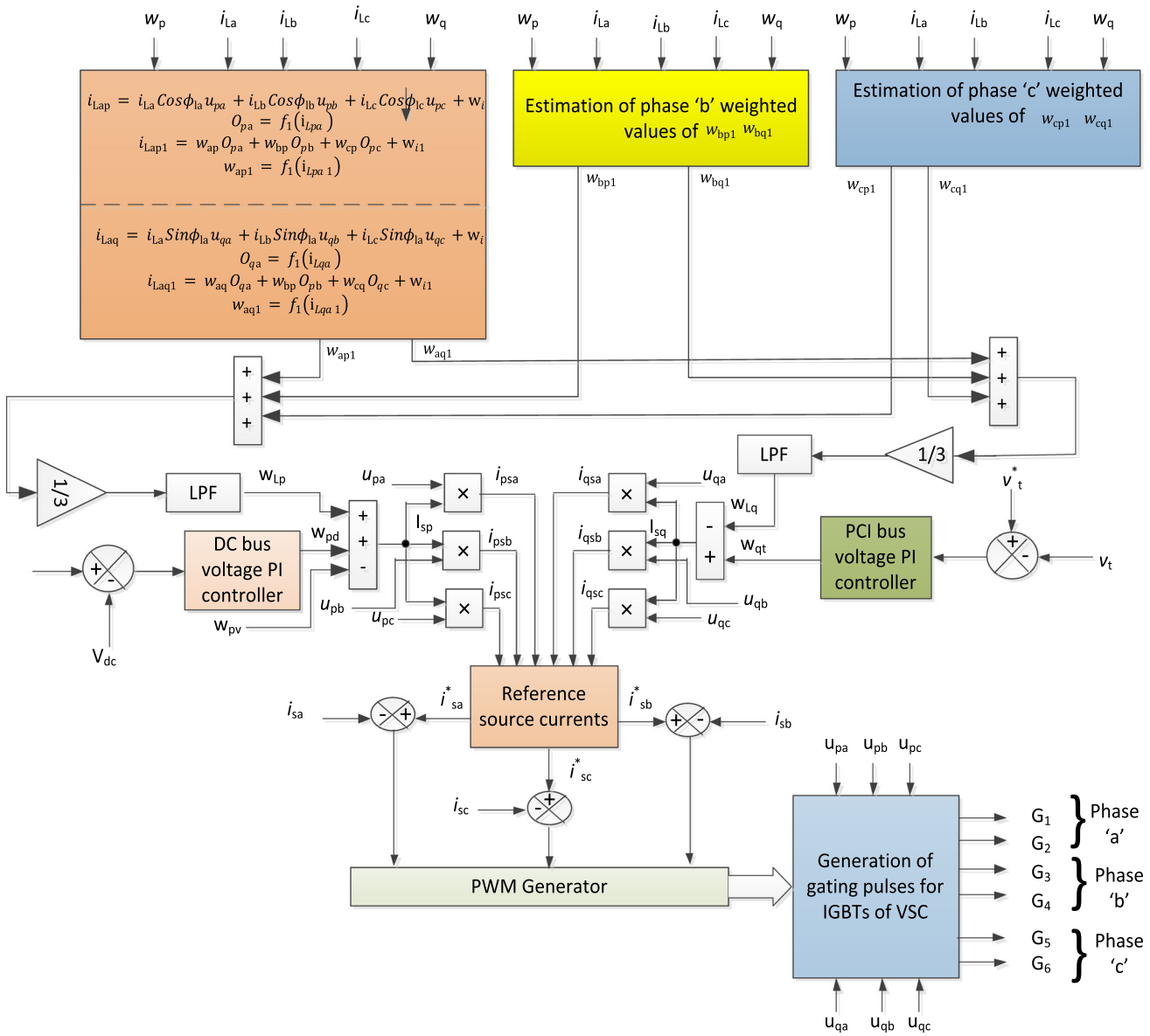


Fig. 3 Proposed control algorithm

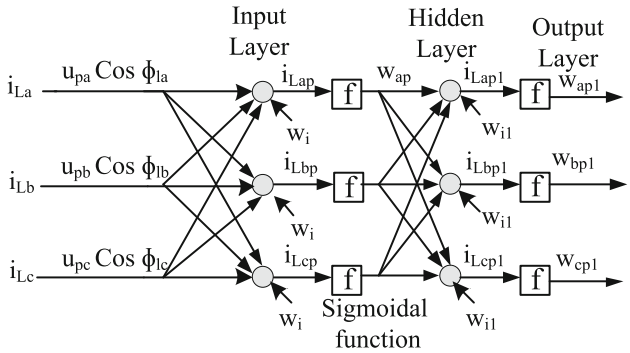


Fig. 4 QNBP NN structure for calculating weighted fundamental active current components

Figure 4 depicts the internal structure of three-layer QNBP neural network which is being used for calculating fundamental active current components.

The extracted components from Eq. (6) are processed with a sigmoid activation function and gives O_{pa}, O_{pb}, O_{pc}:

$$\left.\begin{aligned} O_{pa} &= f_1\left(i_{Lpa}\right) = \frac{1}{1+\mathrm{e}^{-i_{Lpa}}} \\ O_{pb} &= f_1\left(i_{Lpb}\right) = \frac{1}{1+\mathrm{e}^{-i_{Lpb}}} \\ O_{pc} &= f_1\left(i_{Lpc}\right) = \frac{1}{1+\mathrm{e}^{-i_{Lpc}}} \end{aligned}\right\} \tag{20}$$

The output values O_{pa}, O_{pb}, O_{pc} acts as input to hidden layer of the proposed NN with initial weight of w_{i1}. The output of hidden layer can be gives i_{Lap1}, i_{Lbp1} and i_{Lcp1}

with w_{ap}, w_{bp} and w_{cp} as updated weighted values of respective active components.

$$\left.\begin{aligned} i_{Lap1} &= w_{ap}O_{pa} + w_{bp}O_{pb} + w_{cp}O_{pc} + w_{i1} \\ i_{Lbp1} &= w_{ap}O_{pa} + w_{bp}O_{pb} + w_{cp}O_{pc} + w_{i1} \\ i_{Lcp1} &= w_{ap}O_{pa} + w_{bp}O_{pb} + w_{cp}O_{pc} + w_{i1} \end{aligned}\right\} \tag{21}$$

The updated weights values at hidden layer by QNBP given as:

$$\begin{bmatrix} w_{ap} \\ w_{bp} \\ w_{cp} \end{bmatrix} = w_p + \frac{1}{\beta}\left[\left(w_p - w_{ap1}\right)^{-1}\left(w_p - w_{bp1}\right)^{-1}\left(w_p - w_{cp1}\right)^{-1}\right]\begin{bmatrix} O_{pa}^{-1} \\ O_{pb}^{-1} \\ O_{pc}^{-1} \end{bmatrix}\begin{bmatrix} f'\left(i_{Lap1}\right) \\ f'\left(i_{Lbp1}\right) \\ f'\left(i_{Lcp1}\right) \end{bmatrix} \tag{22}$$

Here, w_p depicts the value of average weighted active current component, $f'\left(i_{Lap1}\right)$ which shows first derivative of i_{Lap1}. Also in the above expression, β denotes the learning rate, and it is used to achieve convergence rapidly.

$$\left.\begin{aligned} w_{ap1} &= f_1\left(i_{Lpa1}\right) = \frac{1}{1+\mathrm{e}^{-i_{Lpa1}}} \\ w_{bp1} &= f_1\left(i_{Lpb1}\right) = \frac{1}{1+\mathrm{e}^{-i_{Lpb1}}} \\ w_{cp1} &= f_1\left(i_{Lpc1}\right) = \frac{1}{1+\mathrm{e}^{-i_{Lpc1}}} \end{aligned}\right\} \tag{23}$$

w_{ap1}, w_{bp1} and w_{cp1} are the fundamental weighted values of respective phase currents active components and the mean active component (w_{Lp}) can be calculated as:

$$w_{Lp} = \frac{\left(w_{ap1} + w_{bp1} + w_{cp1}\right)}{3}. \tag{24}$$

Further, fundamental element of reactive current can also be obtained as above given procedure for active component calculation. Quadrature unit templates u_{qa}, u_{qb} and u_{qc} can be taken from Eq. (10).

$$\left.\begin{aligned} i_{Laq} &= i_{La}\mathrm{Sin}\phi_{la}u_{qa} + i_{Lb}\mathrm{Sin}\phi_{la}u_{qb} + i_{Lc}\mathrm{Sin}\phi_{la}u_{qc} + w_i \\ i_{Lbq} &= i_{La}\mathrm{Sin}\phi_{la}u_{qa} + i_{Lb}\mathrm{Sin}\phi_{la}u_{qb} + i_{Lc}\mathrm{Sin}\phi_{la}u_{qc} + w_i \\ i_{Lcq} &= i_{La}\mathrm{Sin}\phi_{la}u_{qa} + i_{Lb}\mathrm{Sin}\phi_{la}u_{qb} + i_{Lc}\mathrm{Sin}\phi_{la}u_{qc} + w_i \end{aligned}\right\} \tag{25}$$

Further, extracted components are processed through a sigmoid activation function to obtain O_{qa}, O_{qb}, O_{qc} (Fig. 5).

$$\left.\begin{aligned} O_{qa} &= f_1\left(i_{Lqa}\right) = \frac{1}{1+\mathrm{e}^{-i_{Lqa}}} \\ O_{qb} &= f_1\left(i_{Lqb}\right) = \frac{1}{1+\mathrm{e}^{-i_{Lqb}}} \\ O_{qc} &= f_1\left(i_{Lqc}\right) = \frac{1}{1+\mathrm{e}^{-i_{Lqc}}} \end{aligned}\right\} \tag{26}$$

Further, these reactive current components processed with these output of sigmoid activation function and feed to

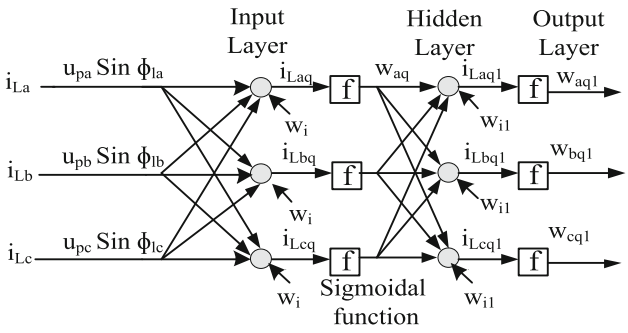


Fig. 5 QNBP NN structure for calculating weighted fundamental reactive current components

hidden layer of neural network as an input considering w_{i1} as initial weight. Moreover, hidden layer output given in Eq. (17) compute the value of fundamental component i_{Laq1}, i_{Lbq1} and i_{Lcq1} with updated weightsw_{aq}, w_{bq} and w_{cq}.

$$\left.\begin{aligned} i_{Laq1} &= w_{aq}O_{qa} + w_{bq}O_{qb} + w_{cq}O_{qc} + w_{i1} \\ i_{Lbq1} &= w_{aq}O_{qa} + w_{bq}O_{qb} + w_{cq}O_{qc} + w_{i1} \\ i_{Lcq1} &= w_{aq}O_{qa} + w_{bq}O_{qb} + w_{cq}O_{qc} + w_{i1} \end{aligned}\right\} \tag{27}$$

The updated weights for hidden layer computed as:

$$\begin{bmatrix} w_{aq} \\ w_{bq} \\ w_{cq} \end{bmatrix} = w_q + \frac{1}{\beta}\left[\left(w_q - w_{aq1}\right)^{-1}\left(w_q - w_{bq1}\right)^{-1}\left(w_q - w_{cq1}\right)^{-1}\right]\begin{bmatrix} O_{qa}^{-1} \\ O_{qb}^{-1} \\ O_{qc}^{-1} \end{bmatrix}\begin{bmatrix} f'\left(i_{Laq1}\right) \\ f'\left(i_{Lbq1}\right) \\ f'\left(i_{Lcq1}\right) \end{bmatrix} \tag{28}$$

In Eq. (18) w_q represent the value of average weighted reactive current element where $f'\left(i_{Laq1}\right)$ shows the very first derivative of i_{Laq1}.

$$\left.\begin{aligned} w_{aq1} &= f_1\left(i_{Lqa1}\right) = \frac{1}{1+\mathrm{e}^{-i_{Lqa1}}} \\ w_{bq1} &= f_1\left(i_{Lqb1}\right) = \frac{1}{1+\mathrm{e}^{-i_{Lqb1}}} \\ w_{cq1} &= f_1\left(i_{Lqc1}\right) = \frac{1}{1+\mathrm{e}^{-i_{Lqc1}}} \end{aligned}\right\} \tag{29}$$

Further, the value of mean reactive component (w_{Lq}) has been obtained by averaging the weighted elements of each phase of the system.

$$w_{Lq} = \frac{\left(w_{aq1} + w_{bq1} + w_{cq1}\right)}{3}. \tag{30}$$

3.2.3 Switching pattern generation for IGBT switches of voltage source converter

This section deals with the generation of reference current and VSC switching signals. Based on the details provided

in above section, the active fundamental element of given load current can be obtained as:

$$I_{Lp} = w_{pd} + w_{Lp} - w_{\mathrm{PV}} \tag{31}$$

Moreover, the fundamental active current component of load is being used to calculate the reference active component of grid current. Further, it is achieved by multiplying the in-phase voltage unit templates and net fundamental active component of current, and can be written as:

$$i_{psa} = I_{Lp} * u_{pa}, \quad i_{psb} = I_{Lp} * u_{pb}, \quad i_{psc} = I_{Lp} * u_{pc} \tag{32}$$

Next, fundamental reactive component of load current can also be calculated likewise the calculation of fundamental active component of current, which is given as:

$$I_{Lq} = w_{qt} + w_{Lq} \tag{33}$$

To calculate, the reference reactive component of the grid side current reference, quadrature voltage unit templates can be multiply with net fundamental reactive component of load current, the same can be obtained as follows:

$$i_{qsa} = I_{Lq} * u_{qa}, \quad i_{qsb} = I_{Lq} * u_{qb}, \quad i_{qsc} = I_{Lq} * u_{qc} \tag{34}$$

Net fundamental grid reference currents $(i^*_{sa}, i^*_{sb}, i^*_{sc})$ given as:

$$i^*_{sa} = i_{psa} + i_{qsa}, \quad i^*_{sb} = i_{psb} + i_{qsb}, \quad i^*_{sc} = i_{psc} + i_{qsc} \tag{35}$$

Further, these calculated reference grid currents $(i^*_{sa}, i^*_{sb}, i^*_{sc})$ is compared with the actual system grid currents to get the switching signals for IGBT switches of VSC. Also, the difference error between above mentioned has been minimized with the help of a PI current regulator.

4 Simulation results

The proposed system is built in MATLAB/Simulink platform. The developed controller supplies the power coming from solar irradiance to the load and utility and further improves power quality. The simulation performance has been analyzed under steady state and dynamic scenarios. Also the performance analysis also done for variable solar irradiances. The performance analysis also carried out under power factor correction (PFC) mode of operation.

4.1 Performance under dynamic linear load

The performance of the developed solar photovoltaic based plants is carried out for dynamic nonlinear load for zero voltage regulation mode and shown in Fig. 6. The parameters observed here are including:

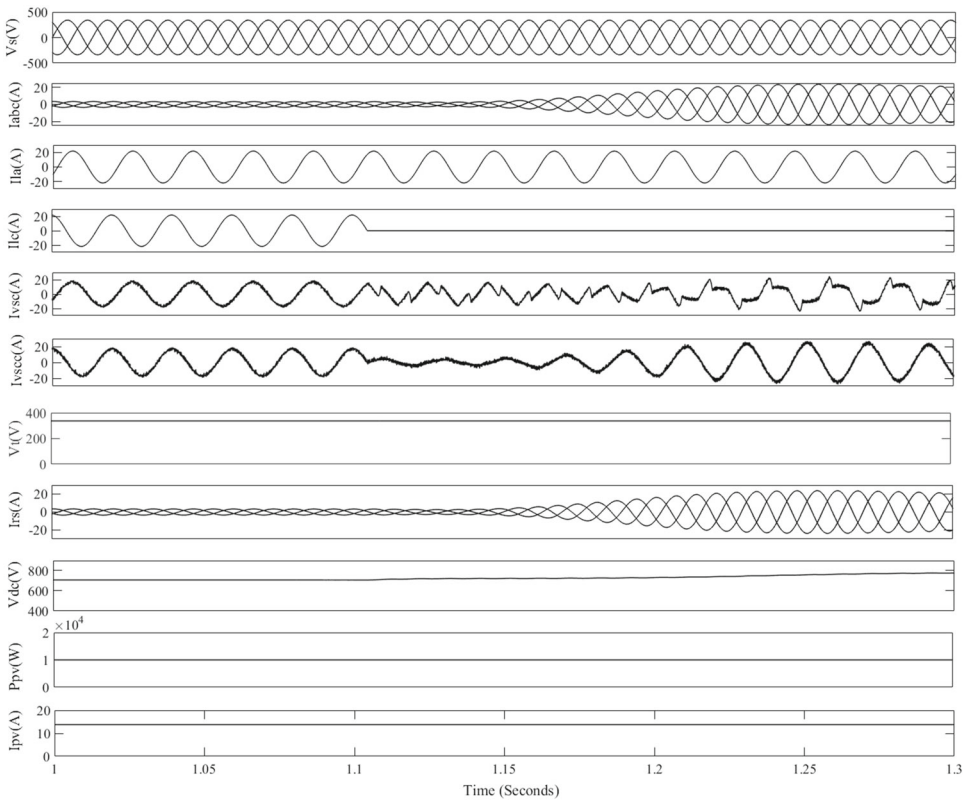


Fig. 6 Performance parameters for dynamic linear load in ZVR mode

Grid voltage (V_s) in volts
Grid current (i_{abc}) in amperes
Load current for phase a and phase c (i_{la} and i_{lc}) in amperes
Inverter current (I_c) in amperes
Terminal voltage (V_t) in volts
DC link voltage (V_{dc}) in volts
Power of solar photovoltaic system (P_{PV}) in watts
Photovoltaic current (I_{PV}) in amperes

The terminal voltage at PCC would change due to drawn of reactive power either by load or from grid.

To create the disturbance in the system, one of the phases like phase c is disconnected at 1.1 s, and the behavior of the grid current is observed and found sinusoidal in nature. This shows the effectiveness of the proposed controller for the grid integrated solar photovoltaic based power plant. Moreover, the values of terminal voltage and the voltage at DC link are also found accurate, and there are no significant fluctuations. In addition, the power factor of the system gets improved and can be observed from the behavior of the grid voltage and current waveforms.

4.2 Performance analysis considering dynamic nonlinear load

Here, the behavior of the system with nonlinear loads in zero voltage regulation mode is evaluated and presented in Fig. 7. Again, the phase c of the system is disconnected at 1.1 s and the parameters like grid current is observed and found sinusoidal in nature. Further, the values of terminal voltage and DC link voltages are also found within the desired limits. From the developed controller, the values of THD at a point of common coupling for the parameters load current, grid current and the grid voltage is 37.48%, 1.04% and 1.05% respectively. Hence, it is evident from Fig. 8a–c that the performance of the systems is good and found in the prescribed limits.

4.3 Performance analysis for variable solar irradiance

The performance of the controller is also validated under varying irradiances considering DC link voltage (V_{dc}) and shown in Fig. 9. A small step change of 20 V is applied in reference DC link voltage (V_{dcref}). It can be clearly seen that with the help of proposed controller, DC link voltage is closely following the reference voltage irrespective of the change in solar irradiance (S). Figure 10 depicts the results of the proposed controller for variations in the solar irradiance. It can be observed that the change in solar irradiance is created at 0.5 s by reducing its value to 600 W/m^2. The proposed controller is effectively maintaining the DC link voltage to its reference and as well as MPP operation is also achieved with system operation at unity power factor (UPF). The response of the proposed system in terms of power is shown in Fig. 11. The solar irradiance is reduced to 500 W/m^2 at 0.5 s and increased back to 1000 W/m^2 at 0.58 s. It can be observed that the PV power is reduced at

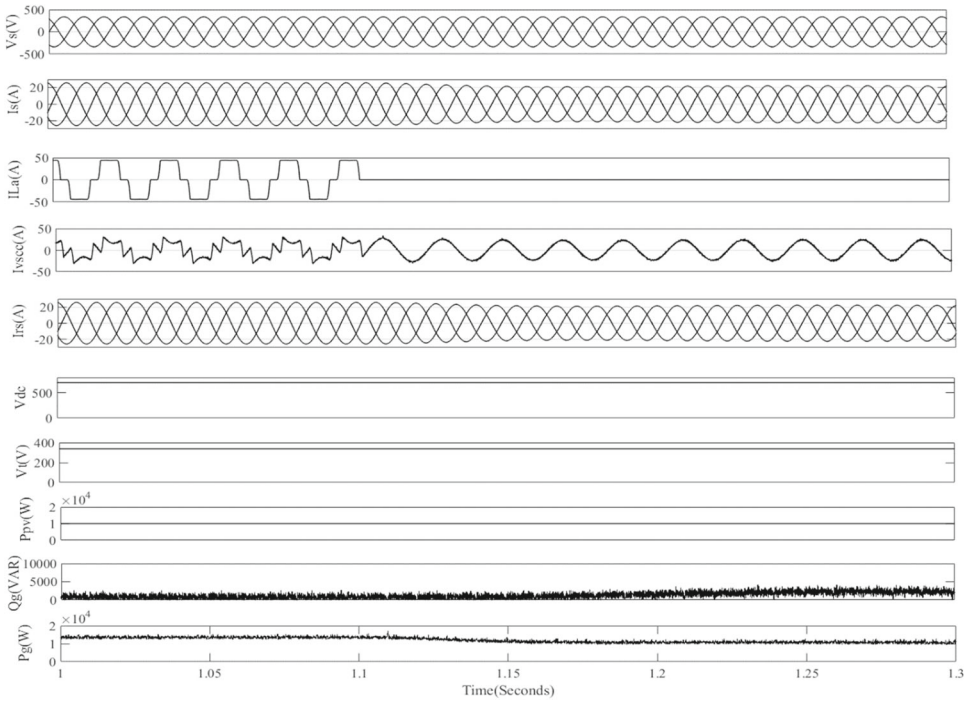


Fig. 7 Performance indices under dynamic nonlinear load

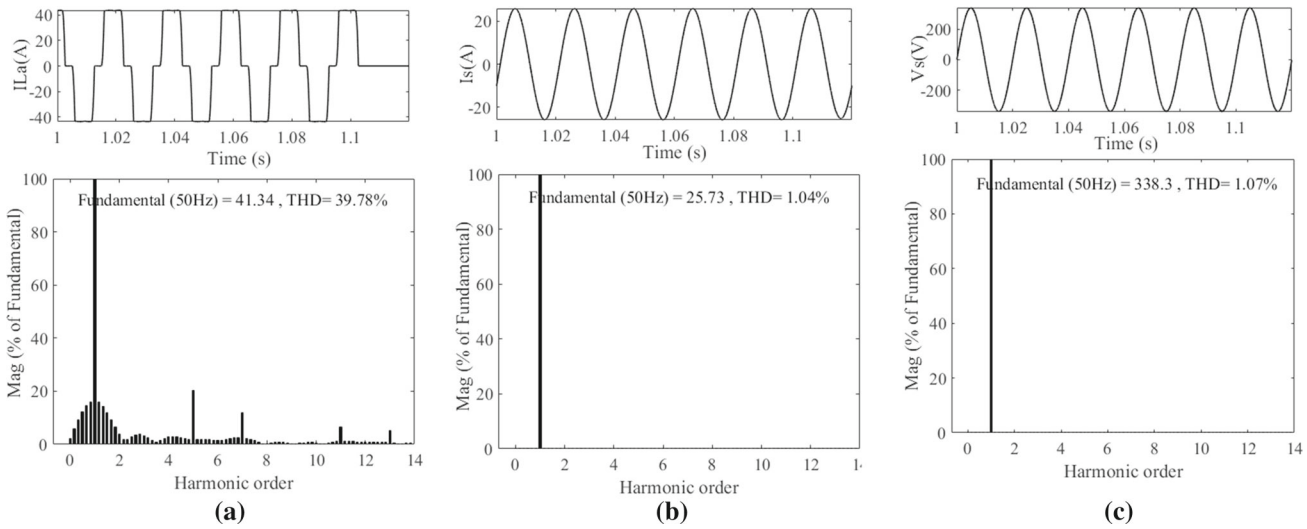


Fig. 8 THD for **a** load current of phase '*a*', **b** grid current and **c** grid voltage in ZVR mode

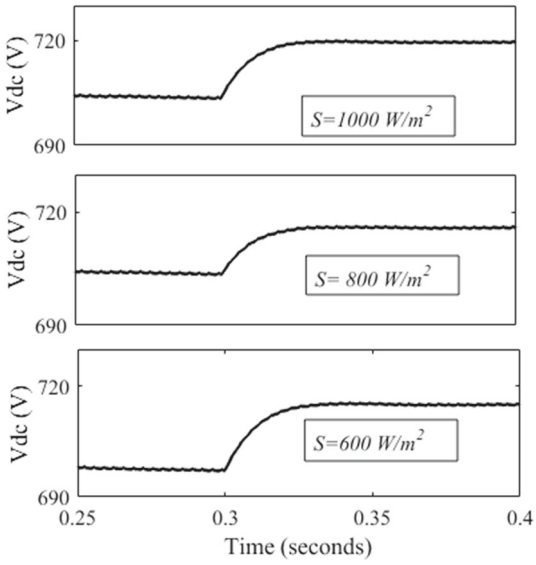


Fig. 9 V_{dc} (V) for various irradiances

0.5 s, hence grid power is increased. The reactive power (Q_g) is constant irrespective change in the solar irradiance.

For comparatively analyzing the same circuit conditions, several most famous proposed techniques are being tested and presented in Table 1. While testing process, nonlinear load is being connected at power control centre (PCC) and grid voltage distortion conditions is considered. The goal of the controller is, to feed the load after load feed and hence, remaining power is transferred to the grid. The obtained THD of grid current is so poor which is achieved by the use of ANN dependent control algorithm. Additionally, THD of grid current is enhanced to 2.78%, by using ANFIS-based control algorithm, whereas the resulted THD of grid current using proposed algorithm is 1.04% only that denotes an important difference, while comparing the art technique's state. The proposed control algorithm is observed better than other conventional, intelligent, and adaptive approaches under dynamic nonlinear load. The given algorithm performs smoother and response found better than other approaches (Fig. 12).

5 Hardware results

The prototype model includes various components development which are presented in detail here. AGV PVE 15 kW solar PV array simulator with $V_{ocmax} = 1500$ V and $I_{scmax} = 30$ A is used as constant DC source to emulate the characteristics of real solar PV array. The design of VSC includes three IGBT modules with a DC voltage storage capacitor. Further, other components are sensing and signal conditioning circuits, AC interfacing inductors, ripple filter (resistance and capacitor banks), linear load and combination of resistance/diode rectifier module as nonlinear loads. Figure 13 depicts the results of practical implementation of the system under power factor correction mode of operation for linear load scenario. Waveforms obtained are: inverter current (I_{vsc}), load current (I_{La}), source side current (I_g) and dc voltage at DC link (V_{dc}). Further, to achieve more clarity about the performance of the proposed controller, other parameters are also shown in Fig. 14. These parameters are given as source side current and source side voltage of phase '*a*'.

Also, the proposed controller also validated for nonlinear load conditions and presented as Fig. 15. Figure 16a shows nonlinear load current and (b) shows the source current and voltage for phase '*a*'. The given load current has THD of 26.5% which is improved to 2.9% for the given algorithm and shown in Fig. 17.

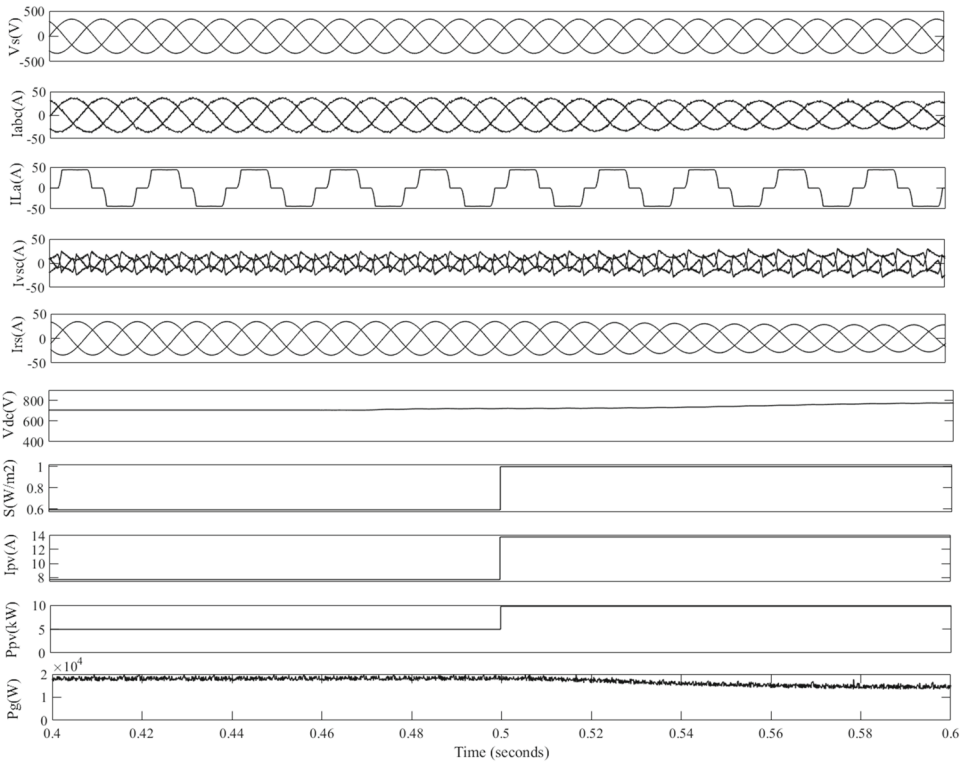


Fig. 10 Performance parameters under nonlinear load conditions for variable solar irradiance

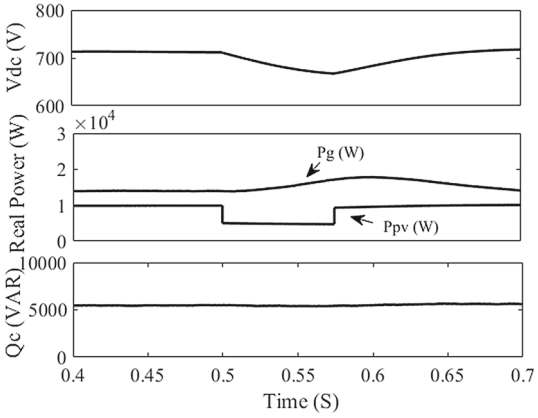


Fig. 11 PV array voltage for MPPT, PV array (P_{PV}) and grid power (P_g), reactive power supplied by inverter (Q_g)

Figure 18 depicts that by removing the loads of phase '*a*', the total load current turns to zero. So, because of the appropriate control operations, grid current is balanced and sinusoidal, and voltage of direct current link is constant. Because of interruption of phase '*a*' load, need of harmonic current is decreased, hence, shape of voltage source converter current is enhanced and tends to sinusoidal, with this nonlinearity of phase '*b*' load current is also decreased. For practical implementation, system used a real grid, hence to create distortions in voltage of grid, inductor has been connected serially to the grid, while another terminal of used inductor is treated as the grid point. The results under the conditions of grid voltage distortion depicts the grid voltage distortion before switching on the inverter, which is an open circuit case. The dynamic impacts on phase '*a*' and phase '*b*' respectively, by removing loads of phase '*a*,' at the time of uneven conditions of load are observed. For validating its efficacy, the controller further tested for variations in solar irradiance and results are presented in Fig. 19. It can be seen from results, that the controller is accurately responding for changes in irradiance.

6 Conclusions

In this research, a proficient controller is designed for the grid supported solar photovoltaic-based power plant. The proposed controller is designed in the MATLAB platform and the simulations are carried out for various disturbances and unbalancing conditions like variation in available solar energy and nonlinear loads. From the results of the proposed controller, it is revealed that the controller is proficient and providing unity power factor mode along with improved power quality features and within the desired

Table 1 Comparative analysis of proposed controller with existing intelligent controllers

Algorithm	Settling time (s)	DC link voltage overshoot (V)	DC link voltage undershoot (V)	THD % (Grid current)
SRF	0.096	820	763	5.63
I cos ϕ	0.095	850	786	4.16
Conductance fryze	0.11	957	796	5.79
ANN based	0.095	815	774	3.25
ANFIS based	0.094	820	779	2.78
FLC based	0.095	815	780	2.25
Adaptive	0.095	810	770	2.18
I cos ϕ QNBP NN	0.091	790	775	1.04

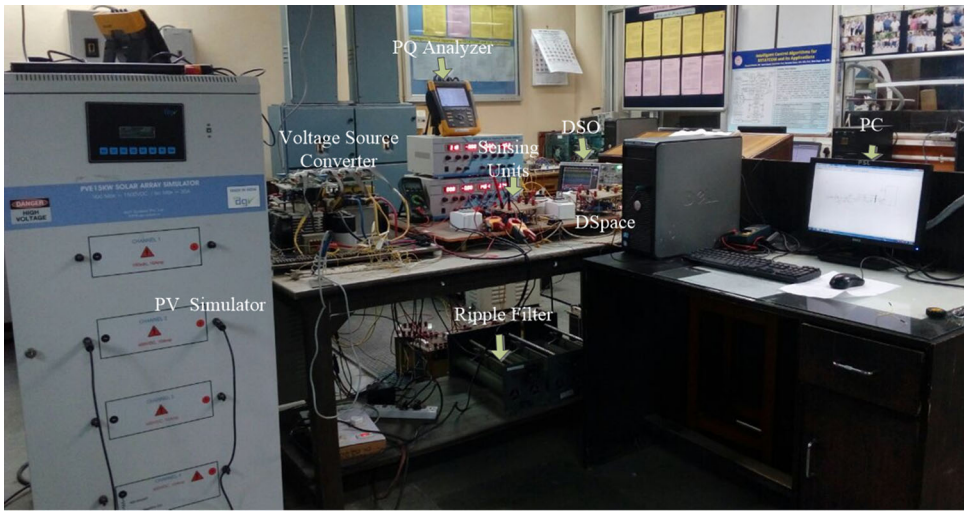


Fig. 12 Laboratory setup

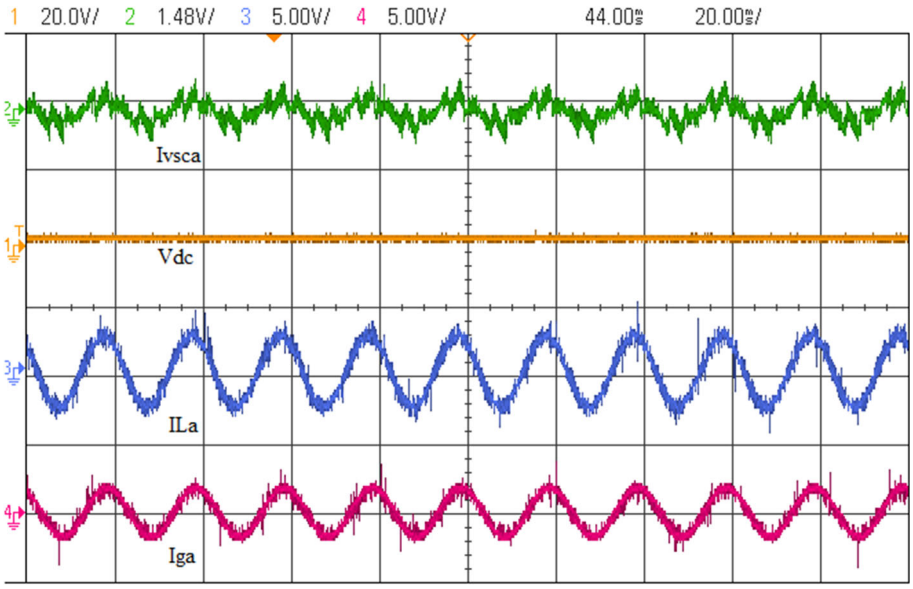
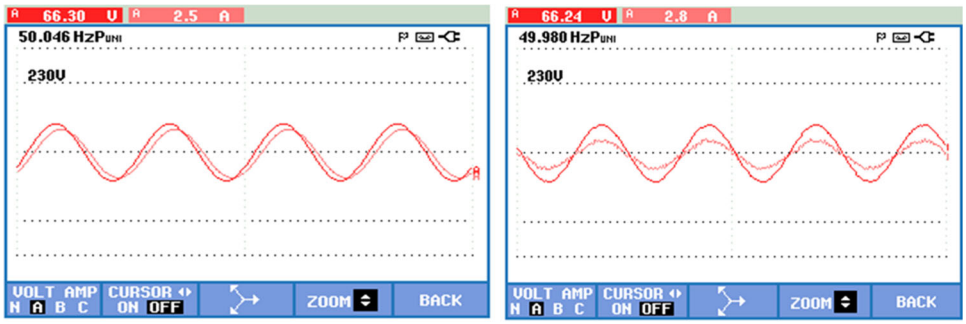


Fig. 13 Controller performance for linear load



(a) (b)

Fig. 14 **a** Grid voltage and current of phase '*a*' and **b** grid voltage and current of phase '*a*' for PFC mode

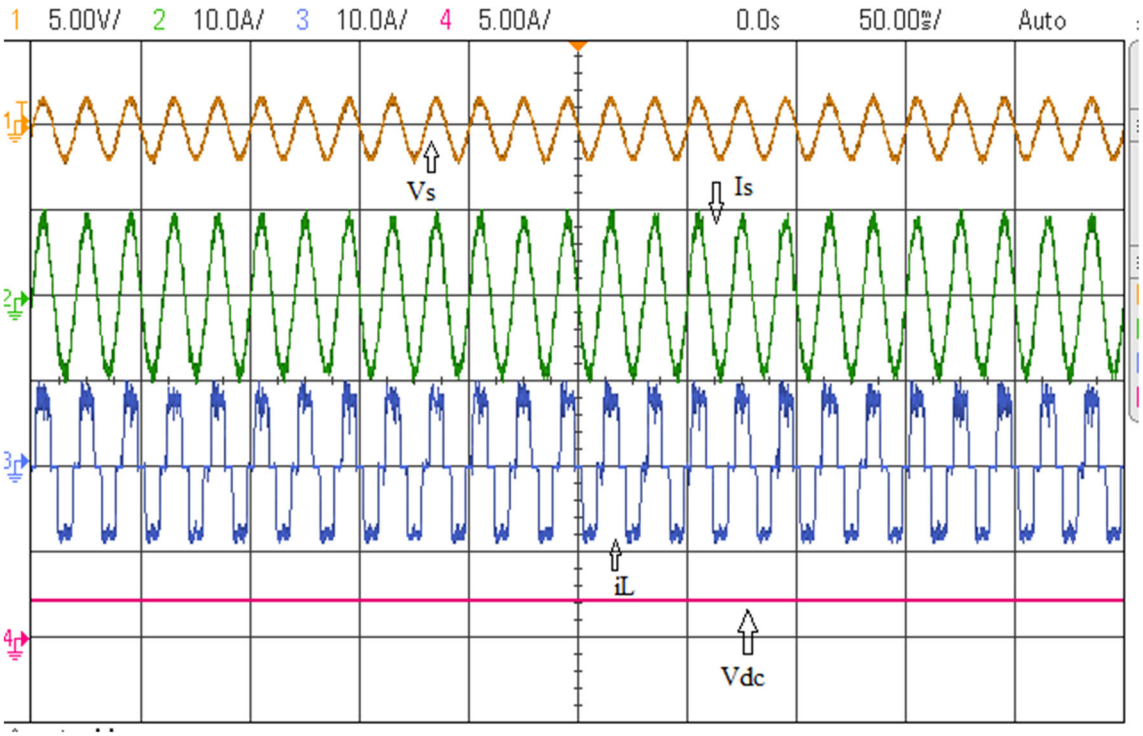
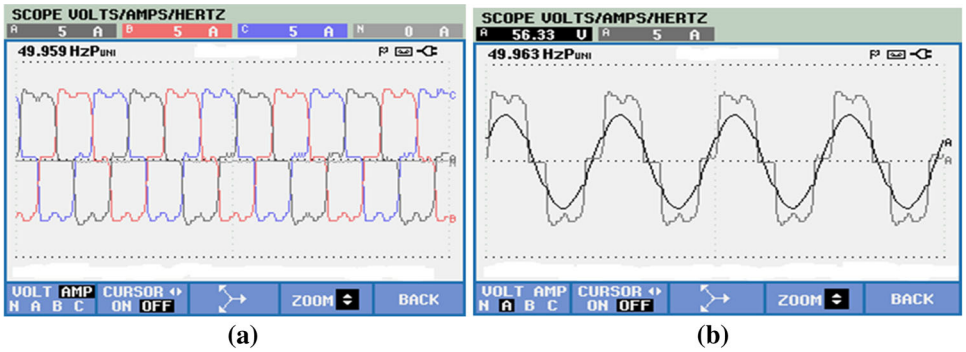
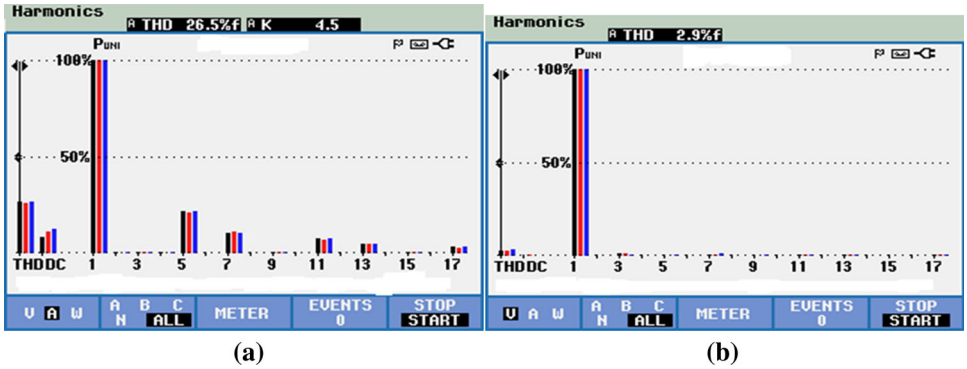


Fig. 15 Performance parameters under nonlinear load conditions



(a) (b)

Fig. 16 **a** Nonlinear current and **b** current and voltage of phase '*a*'



(a) (b)

Fig. 17 **a** THD spectrum of nonlinear load and **b** improved THD for grid current under nonlinear load

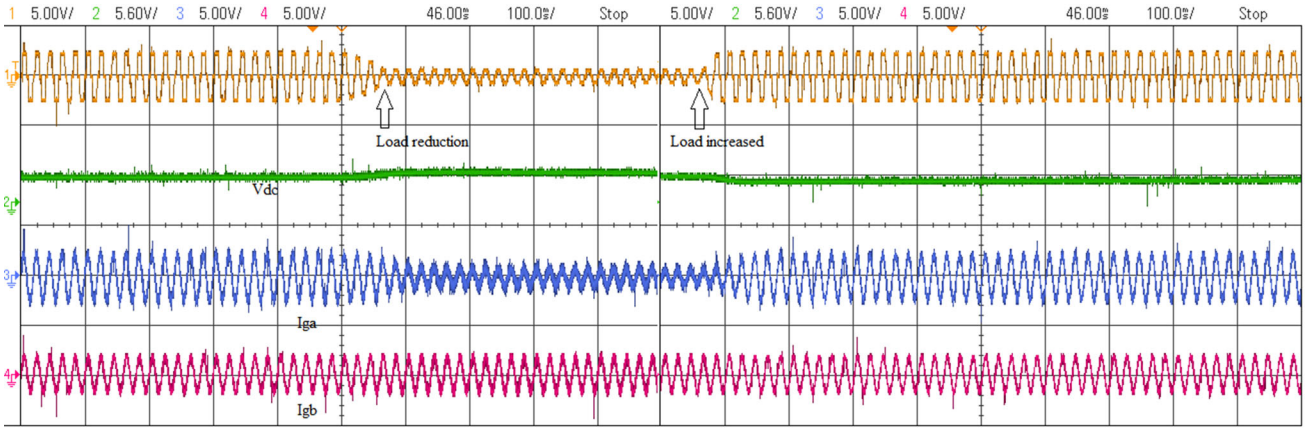


Fig. 18 Performance for dynamic nonlinear load

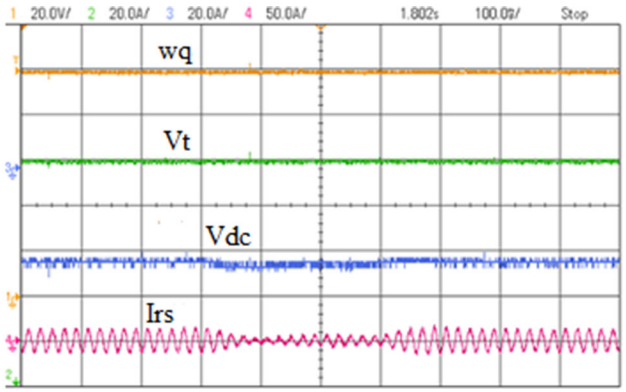


Fig. 19 Performance for variable solar irradiance

limits and IEEE-519 standard. To validate the system performance a laboratory prototype is also developed and simulation results are verified. Moreover, the reactive power compensation along with elimination of harmonics is achieved at unity power factor.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval The article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

Agarwal RK, Hussain I, Singh B (2017) Three-phase single-stage grid tied solar PV ECS using PLL-less fast CTF control technique. IET Power Electron 10(2):178–188

Arya SR, Singh B (2014) Neural network based conductance estimation control algorithm for shunt compensation. IEEE Trans Ind Inform 10(1):569–577

Arya SR, Singh B, Chandra A, Al-Haddad K (2014) Learning-based anti-Hebbian algorithm for control of distribution static compensator. IEEE Trans Ind Electron 61(11):6004–6012

Arya SR, Niwas R, Bhalla KK, Singh B, Chandra A, Al-Haddad K (2015) Power quality improvement in isolated distributed power generating system using DSTATCOM. IEEE Trans Ind Appl 51(6):4766–4774

Chaudhary P, Rizwan M (2015) A Grid synchronization method based on adaptive notch filter for PV system with modified MPPT. World Acad Sci Eng Technol 9(7):1162–1167

Chaudhary P, Rizwan M (2017) A predictive current control for solar PV fed VSI in distribution system. In: 2017 IEEE international conference on environment and electrical engineering and 2017 IEEE industrial and commercial power systems Europe (EEEIC/I&CPS Europe), Milan, pp 1–7

Chaudhary P, Rizwan M (2018a) Voltage regulation mitigation techniques in distribution system with high PV penetration: a review. Renew Sustain Energy Rev 82:3279–3287

Chaudhary P, Rizwan M (2018b) Energy management supporting high penetration of solar photovoltaic generation for smart grid using solar forecasts and pumped hydro storage system. Renew Energy 118:928–946

Chaudhary P, Rizwan M (2018c) Hybrid control approach for PV/FC fed voltage source converter tied to grid. Int J Hydrog Energy 43:6851–6866

Chaudhary P, Rizwan M (2019) Intelligent approach-based hybrid control algorithm for integration of solar photovoltaic system in smart grid environment. IET Smart Grid 2(3):445–454

Chen YK, Yang CH, Wu YC (2009) Robust fuzzy controlled photovoltaic power inverter with Taguchi method. IEEE Trans Aerosp Electron Syst 38(3):940–954

Hamid MI, Jusoh A, Anwari M (2014) Photovoltaic plant with reduced output current harmonics using generation-side active power conditioner. IET Renew Power Gener 8(7):817–826

He J, Li YW (2013) Hybrid voltage and current control approach for dg-grid interfacing converters with LCL filters. IEEE Trans Ind Electron 60(5):1797–1809

IEEE Recommended practices and requirement for harmonic control on electric power system. IEEE Std. 519 (1992)

Jain C, Singh B (2016) An adjustable DC link voltage based control of multifunctional grid interfaced solar PV system. IEEE J Emerg Sel Top Power Electron 99:1

Janpong S, Areerak KL, Areerak KN (2011) A literature survey of neural network applications for shunt active power filters. World Acad Sci Eng Technol 60:392–398

Jayachandran J, Sachithanandam R (2015) Neural network based control algorithm for DSTATCOM under nonideal source voltage and varying load conditions. Can J Electr Comput Eng 38(4):307–317

Kannan VK, Rengarajan N (2014) Investigating the performance of photovoltaic based DSTATCOM using $I\ cos\ \phi$ algorithm. Electron Power Energy Syst 54:376–386

Karimi M, Mokhlis H, Naidu K, Uddin S, Bakar AHA (2016) Photovoltaic penetration issues and impacts in distribution network: a review. Renew Sust Energy Rev 53:594–605

Kulkarni RV, Venayagamoorthy GK (2009) Generalized neuron: feedforward and recurrent architecture. Elsevier Neural Netw 22:1011–1017

Kumar AP, Mangaraj M (2013) DSTATCOM employing hybrid neural network control technique for power quality improvement. IET Power Electron 10(4):480–489

Kumar C, Mishra MK (2014) A multifunctional DSTATCOM operating under stiff source. IEEE Trans Ind Electron 61(7):3131–3136

Mahmud MA, Pota HR, Hossain MJ (2014) Nonlinear current control scheme for a single-phase grid-connected photovoltaic system. IEEE Trans Sustain Energy 5(1):218–227

Mangaraj M, Panda AK (2017) Performance analysis of DSTATCOM employing various control algorithms. IET Gener Transm Distrib 11(10):2643–2653

Masoum MAS, Dehbonei H, Fuchs EF (2002) Theoretical and experimental analysis of photovoltaic systems with voltage and current based maximum power-point tracking. IEEE Trans Energy Convert 7(4):514–522

Mazumdar J, Harley RG (2008) Recurrent neural networks trained with back propagation through time algorithm to estimate nonlinear load harmonic currents. IEEE Trans Ind Electron 55(9):3484–3491

Pandiarajan N, Ramaprabha R, Ranganath M (2011) Application of circuit model for photovoltaic energy conversion system. Int J Adv Eng Technol 2(4):118–127

Rajput MA, Beniwal N, Pradhan S, Singh B, Mishra S (2017) Single-phase grid interactive solar PV system having active shunt capabilities. In: 2017 7th international conference on power systems (ICPS), Pune, India, pp 194–199

Reisi AR, Moradi MH, Jamsab S (2013) Classification and comparison of maximum power point tracking techniques for photovoltaic system: a review. Renew Sustain Energy Rev 19:433–443

Rizwan M, Chaudhary P, Ahmad T (2014) Performance analysis of maximum power point tracking techniques for photovoltaic systems. Adv Sci Lett 20:1231–1247

Sing B, Jain C, Goel S (2014) ILST control algorithm of single-stage dual purpose grid connected solar PV system. IEEE Trans Power Electron 29(10):5347–5357

Singh B, Arya SR (2014) Back-propagation control algorithm for power quality improvement using DSTATCOM. IEEE Trans Ind Electron 61(3):1204–1212

Singh B, Jayaprakash P, Kumar S, Kothari DP (2011) Implementation of neural network controlled three leg VSC and a transformer as three phase four wire DSTATCOM. IEEE Trans Ind Appl 47(4):1892–1901

Singh B, Jayaprakash P, Kothari DP, Chandra A, Al Haddad K (2014) Comprehensive study of DSTATCOM configurations. IEEE Trans Ind Inform 10(2):854–870

Singh B, Jain C, Goel S, Chandra A, Al-Haddad K (2016) A multifunctional grid-tied solar energy conversion system with ANF-based control approach. IEEE Trans Ind Appl 52(5):3663–3672

Varma RK, Rahman SA, Vanderheide T (2015) New control of PV solar farm as STATCOM (PV-STATCOM) for increasing grid power transmission limits during night and day. IEEE Trans Power Deliv 30(2):755–763

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/344191988

Denoising SMS Text Using Ternary Tree for FAQ Retrieval

Article *in* International Journal of Advanced Science and Technology · January 2020

CITATIONS
0

READ
1

2 authors, including:



Manoj Kumar
Delhi Technological University
25 PUBLICATIONS **33** CITATIONS

SEE PROFILE

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

SMS-based FAQ Retrieval View project

All content following this page was uploaded by Manoj Kumar on 10 September 2020.

The user has requested enhancement of the downloaded file.

International Journal of Advanced Science and Technology
Vol. 29, No. 3, (2020), pp. 13200 – 13209

Denoising SMS Text Using Ternary Tree for FAQ Retrieval

Manoj Kumar[1], Vipra Bhatt[2]
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
[1]mkumarg@dce.ac.in, [2]viprabhatt81@gmail.com

Abstract

With the increase in mobile technology, everyone is using mobile phones for various purposes, from only making calls to using it for accessing the internet, gaming, etc. Internet service is not available on every mobile phone, but the facility of Short Messaging Services (SMS) is on every phone. So, many people choose SMS for information retrieval. The SMS can be used to get a quick response in various fields such as healthcare, railways, insurance, banking and travel. But the SMS texts are made noisy by intentional and non-intentional errors, and there is a need to find the correct word to substitute them. This paper proposed efficiently denoising the SMS term by finding matching prefix and suffix length using a ternary search tree and calculating similarity score between the noisy token and the correct English words using Longest Common Subsequence (LCS). We see that the ternary tree is space-efficient compared to the trie data structure used in the previous approach to efficiently denoise the SMS token. It also helps in completing the task in less time compared to using a hash table. We compute the overall score by combining the prefix matching length, suffix matching length, and similarity score. We experimented using some noisy SMS words generally used by the people and compare our model with some previous models. Our model outperforms all of these earlier approaches.

Keywords: *Frequently asked questions (FAQs), Information Retrieval, Noisy SMS, SMS query, Ternary Search Tree (TST).*

1. Introduction

With the increase in mobile technology, everyone is using mobile phones nowadays. A lot of information is available over the internet and is freely available to us. But in our country, not every individual uses the internet facility as most of the population doesn't know how to use the smartphones, and they use the lost cost feature phone. The facility of SMS is available on every mobile phone. So, everyone can use SMS for accessing information from the frequently asked questions (FAQs) in various fields like agriculture, health, travel, railway reservation, insurance, banking, etc. Using SMS for information retrieval saves time and money. But, the number of characters is limited in SMS, i.e., 160 per SMS, so there is a need to shorten the word while writing the SMS query, which makes the SMS noisy. Also, there are many other factors such as the small screen, damaged display, and multi-tap keyboard in low-cost mobile phones where the same key is pressed repeatedly to type a character, QWERTY keyboard, which makes SMS query noisy.

ISSN: 2005-4238 IJAST
Copyright ⓒ 2020 SERSC

Such noisy texts are the combination of unintentional errors and a few intentional dropping of characters, phonetic substitutions, and abbreviations. Different types of noise [1, 2] include the deletion of vowels and repeated characters to make the word shorter (such as "ltr" for "letter" or "hpy" for "happy"). Truncation of words (such as "exam" for "examination" or "univ" for "university") is done to reduce the number of characters in the word. The use of abbreviations (such as "OMG" for "Oh My God", "LOL" for "Laughing Out Loud", or "BRB" for "Be Right Back") are widespread amongst the youngsters. Phonetic substitution, i.e., the same sounding words are replaced by their shorter substitute (such as "tym" for "time", "4" for "for" or "four" and "gr8" for "great"). Deleting articles and pronouns from the sentences (such as "gng mkt" for "I am going to market") to reduce its length. In dialects and informal use, many words are combined to a single word (such as "lemme" for "let me" or "wanna" for "want to"). Other types of noise include grammatical mistakes, spelling errors, mixing of words of two languages such as the use of interjections ("ooch", "ooh," "huh") and emoticons (☺, ☹), and many others are an example of noise in texts. The use of Hinglish while speaking and writing texts (for example, "Hungry Kya?" - the tagline of Domino's pizza is the combination of an English and a Hindi word) is another type of noise. These words may look normal for humans to understand but are difficult for a computer or an automated system to understand. So, there is a need to denoise them.

SMS based FAQ access can be done using a Human Intervention or by an automated FAQ retrieval system. In the human intervention based method, a human expert is required to interpret the SMS query the user has asked and return the answer to the user. In an automated SMS based FAQ retrieval system, the query is asked via SMS, and the experts predesign the FAQ corpus. The task is to find the question from the FAQ corpus that matches the most with the SMS, and then its corresponding answer is returned to the user. Many models have been proposed to automate the SMS based FAQ retrieval system. Our paper mainly focuses on the algorithm used in the paper by Agarwal et al. [3] to denoise the SMS text and uses the ternary tree to calculate matching prefix and suffix length. We try to remove noise from the SMS words efficiently by finding the top seven matching candidates for the noisy word. The overall score to calculate the matching depends on matching prefix length and matching suffix length and similarity score between the SMS words and the English dictionary words. Based on the total score, we generate a list of the top seven best matching candidates for each of the noisy token.

The rest of the paper is organized as follows. Section 2 describes the prior work done in this area. Terminologies used is discussed in section 3. In section 4, we describe our proposed approach, the various changes we made in the previous model. In section 5, we provide the implementation details and the experimental results. Finally, we concluded the paper in section 5.

2. Prior work

A significant amount of work [3-10] has been done in the field of denoising SMS queries. The text written in SMS language is generally misspelled, transliterated, have grammatical mistakes, abbreviations, etc. Kothari et al. [4] were the first to develop a system to handle this. They used the combinatorial search approach and handled semantic variation along with the similarity using the synonym dictionary. They use two algorithms, namely, naïve and prune algorithms, to generate a list of the candidate terms for the noisy tokens. They formulated a dictionary from all the words appearing in the FAQ set and then define similarity for all words of the SMS query to the dictionary terms. The similarity is determined

ISSN: 2005-4238 IJAST
Copyright ⓒ 2020 SERSC

by the longest common subsequence ratio to the edit distance between noisy SMS token and the dictionary word and is given as:

$$\alpha(t,s) = \begin{cases} \frac{LCSRatio(t,s)}{EditDistance_{SMS}(t,s)}, if\ t[0] == s[0] \\ 0 \qquad\qquad\qquad\quad ,otherwise \end{cases} \tag{1}$$

where LCS ratio is the ratio of the length of LCS to the length of the dictionary term and edit distance is the number of operations (substitutions, deletions, and insertions) required to change one string to another.

The paper given by D. Shaikh et al. [6] uses the same method in [4] to calculate the similarity measure. They make the additional change in the scoring function from [4], adding length score and proximity score along with the similarity score. They give their model for monolingual English and monolingual Hindi tasks and achieved excellent results. The list generation and candidate set generation remains the same as the previous method [4].

The method used by Shivhre et al. [8] for finding the ranked list of words in the same as defined in [7]. They use the weight function to find the similarity between the dictionary terms and the noisy SMS token. The weight function between the SMS token and the dictionary term uses the LCS ratio between them, the similarity ratio (defined using Ratcliff/ Obershelp algorithm), the inverse document frequency of the word in the dictionary, and the Levenshtein distance between the two. Similarity ratio is given twice the number of common characters between SMS and dictionary terms to the number of characters in both. These methods, combined, give the measure to find the weight for all dictionary term w.r.t. The noisy SMS token. Weight of the dictionary term t w.r.t. SMS token s is given by:

$$Weight(t,s) = \frac{LCSR(t,s) * SMRatio(t,s) * IDF(t)}{LevDistance(t,s)} \tag{2}$$

S. Mhaisale et al. [9] use the weighted edit distance (WED) to find the similarity between the SMS and the dictionary word. Instead of giving equal weights to insertion, deletion, and substitution as in constant edit distance. Mhaisale give a relative importance to each of them with substitution being the most frequent operation, followed by insertion and deletion being the least. They give less weightage to more frequent operations. Also, the insertion of vowels and insertion of trailing characters are the most frequent operations, so, are given less weightage. The weighted edit distance between SMS token s and the dictionary term t is given as:

$$WED(t,s) = 0.3 * x_1 + 0.5 * x_2 + 0.3 * x_3 + 0.5 * x_4 + 2.0 * x_5 \tag{3}$$

Where x_1, x_2, x_3, x_4, x_5 denotes the number of vowel insertions, consonant insertions, insertions at the end of token, vowel substitutions, and deletions, respectively.

The similarity in their approach was based on inverse document frequency, LCS ratio, and the weighted edit distance of the dictionary term t and noisy term s and it is given as:

$$Sim(t,s) = \frac{IDF(t) * LCSRatio(t,s)}{WeightedEditDistance(t,s)} \tag{4}$$

As to prune the search space, R. Batra et al. [10] limits the search space from [4], by matching SMS token with the dictionary words whose length is greater or equal to the SMS token which results in improving the computation time and hence improves the result. Also, because the user may type additional characters or may commit typos, the relaxation of two

ISSN: 2005-4238 IJAST
Copyright ⓒ 2020 SERSC

characters is given. According to them, consider only those words that satisfy the given condition: LCS(s,t) >= length(s) – 2. Also, if the LCS of the SMS token and the dictionary term is equal to the length of the SMS word, then the similarity score is increased by one, i.e., $if\ LCS(s,t) == strlen(s), then\ SM = SM + 1$.

Agarwal et al. [3] denoises the SMS token by calculating a similarity measure based on prefix matching, and suffix matching that was computed using prefix and suffix tries respectively and using the similarity ratio. Prefix trie was made to store all the dictionary words, and the noisy word is then searched in the trie to calculated matching prefix length for all the dictionary terms. Similarly, suffix matching is done by reverse storing the dictionary terms. Then, use all the three factors together and calculate an overall score to find the substitute for the noisy SMS token. The overall score between dictionary term t and SMS tokens s is given as:

$$Overall_{Score}(t) = 0.25 * Prefix_{Score}(t,s) + 0.25 * Suffix_{Score}(t,s) + Sim_{Score}(t,s) \quad (5)$$

To reduce the memory space, we use a ternary search tree to calculate the length of prefix and suffix matching. Then we find the similarity of every noisy word with all the dictionary words and compute their respective similarity score. Combining all the three factors: prefix score, suffix score, and similarity score, we efficiently denoises the SMS token.

3. Terminologies used

The noisy SMS query hinders the automated FAQ retrieval system. So, there is a need to efficiently denoise the SMS text. Denoising of SMS text requires finding the closeness of the SMS token with all the dictionary terms. A ternary search tree is used to find the length of prefix and suffix matching and Longest Common Subsequence to find a similarity score.

3.1. Ternary search tree (TST)

A ternary search tree is a type of trie data structure that uses 3 pointers instead of 26 (low memory space). Every node of the ternary tree has 3 pointers (left, middle and right) along with the EndOfword bit and the value field where the character is stored. A node represents each character of a string. If the character is less than the current node, then it goes to the left node. Similarly, if the character is greater than the current node, it goes to the right. And if it is equal to the current node, it goes to the middle one. It is used for various applications such as for storing string prefix, for autocomplete suggestions, etc. The average time complexity for traversal, insertion, and deletion is O ($\log_3$ n), where n is the height of the TST.

Table 1.Time Complexities of TST

	Average	**Worst**
Lookup	O ($\log_3$ n)	O (n)
Insert	O ($\log_3$ n)	O (n)
Delete	O ($\log_3$ n)	O (n)

Ternary tree combines the advantage of less space of the binary tree and little character search time for the tries. Ternary tree search is faster than hashing for the words not present in TST.

ISSN: 2005-4238 IJAST
Copyright ⓒ 2020 SERSC

3.2. Longest common subsequence (LCS)

LCS is given by the number of common characters in the two string. If SMS token s is of length i and dictionary term t is of length j, then LCS is given by:

$$LCS\ [i][j] = \begin{cases} 0 & ,if\ i = 0\ or\ j = 0 \\ a + 1 & ,if\ s[i-1] = t[j-1] \\ max\ (b, c) & ,otherwise \end{cases} \quad (6)$$

Where $a = LCS[i-1][j-1]$ (7)

$b = LCS[i-1][j]$ (8)

$c = LCS[i][j-1]$ (9)

The higher the value of the longest common subsequence, the higher is the matching of the SMS token and the dictionary word.

4. Proposed method

To develop an automated FAQ retrieval model, we need to remove noise from the SMS query. To get the correct word for the noisy SMS token, we consider three factors: prefix, suffix, and similarity score. Prefix and suffix matching are done using a ternary search tree. We use a ternary search tree (TST) to store all the words of the English dictionary. Each node represents a single character and points to three children- left, middle, and right. TST is used to calculate the matching prefix length that helps to find the closeness of the noisy SMS token with the dictionary terms. Similarly, to find the matching suffix length, we store all the words of the English dictionary in the reverse order in TST and similarly calculate the closeness as the prefix. The limitation of our model is that it works well for the single noisy token but fails to denoise the two or more terms abbreviated together. For example, abbreviations such as "BRB" for "Be Right back" are not denoised by our model.

4.1. Prefix matching (Prefix score)

Prefix ternary tree is created to store all the words of the dictionary. The noisy SMS token is then matched with all the dictionary terms. For the terms having minimum prefix matching of two (i.e., if two or more starting characters are matched) is stored in a hash table with key as dictionary term and the value as the prefix length score. It takes into account the noisy words that are truncated, i.e., the words whose trailing characters are cut off. For example, "nov" for "november". Here, the prefix score for the word "november" is 3. These words are correct in the beginning, so finding the prefix score helps us to get the correct alternative for such noisy words.

4.2. Suffix matching (Suffix score)

For suffix matching, we store all the dictionary terms in the reverse order and reverse the noisy SMS token. Then this reversed word is matched with all the words in the suffix ternary tree, and the words with matching suffix are stored in the hash with key as dictionary term and value as the matching suffix length. The suffix ternary tree is created in the reverse order of the prefix ternary tree to reduce the time complexity as finding suffix length by directly

ISSN: 2005-4238 IJAST
Copyright ⓒ 2020 SERSC

creating a suffix tree (without reversing the dictionary terms) would be very complicated and time-consuming. Minimum matching of length one is considered in case of suffix matching.

4.3. Similarity Score

The third factor is the similarity score that takes into account the common characters between the noisy term and the dictionary term. The noisy word is matched with all the dictionary terms, and the similarity score is calculated for all terms present in the dictionary. The concept of the LCS is used to find the similarity score between the noisy token and the dictionary term. The similarity of SMS token s and the dictionary term t is given as:

$$Sim(t,s) = \frac{2 * LCS\,(t,s)}{length(t) + length(s)} \tag{10}$$

Further, If the first character of noisy token and dictionary term is matched, we increase the similarity score by 0.1 times of Sim (t,s). Therefore, the similarity score is given as:

$$Sim_{Score}(t,s) = \begin{cases} Sim(t,s), & t[0] != s[0] \\ 1.1 * Sim(t,s), & ,otherwise \end{cases} \tag{11}$$

4.4. Overall Score

We calculated the overall score for each word of the top 7 terms in the prefix score list, suffix score list, and similarity score list. Prefix score of 0 is given for the term not present in the prefix-list but is present in any of the other two to calculate the overall score. Similar to the suffix score and similarity score. To obtain the overall score for each word of the dictionary corresponding to a noisy SMS word, we combine the above three score calculated as:

$$Overall_{Score}(t) = w_1 * Prefix_{Score}(t,s) + w_2 * Suffix_{Score}(t,s) + w_3 * Sim_{Score}(t,s) \tag{12}$$

Where t is the dictionary term, s is the SMS token, and w_1, w_2, and w_3 are the constant weights assigned to prefix score, suffix score, and similarity score and are given values of 0.25, 0.2, and 1, respectively. The values of these weights are assigned using the experimental results that is discussed in the next section.

More weightage is given to the similarity score because the similarity score considers the ratio, while the prefix and suffix score takes the number of characters matched into consideration. We sort the entries by value (score) from the hash table for the overall score and choose the top 7 results in the output list.

5. Experiments and results

To verify the performance of our proposed model, we test it on Mieliestronk's words list [11] as the dictionary and compare it with Agarwal et al.'s [3] and Kothari et al.'s [4] method. Mieliestronk's words list is comprised of 58000+ English words, and we make 110 noisy words to test our system. To assess quantitatively, we are considering Accuracy and MRR (Mean Reciprocal Rank) to compare our model's output with the previous ones. Accuracy is given by the number of correct output retrieved to the total number of output retrieved. MRR is used as another evaluation measure and is given by:

ISSN: 2005-4238 IJAST
Copyright ⓒ 2020 SERSC

$$MRR = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{rank_i} \quad (13)$$

Where 1/rank is the reciprocal rank. Rank = 1 is given when the correct word is found at the top of the list. Rank = 7 for the word at 7th position in the list. Reciprocal rank is given the value 1/8 for the words not found correct in the top 7 list because the output will be at some later position in the list. We use MRR because it is easy to compute and give focus on the first relevant item of the list; thus, it helps in finding the best match.

There are 3 main parameters in our novel method, i.e., weight for prefix score w_1, weight for suffix score w_2, and weight for similarity score w_3. To study these parameters, we carried out different experiments, where we had kept one parameter variable and the other two as fixed. The results of the experiments are displayed in table 2-4.

Table 2. Parameter w_1 study

Metric	**w_1, w_2, = 0.2, w_3 = 1.1**				
	0.15	**0.20**	**0.25**	**0.30**	**0.35**
Accuracy	90.740	91.666	**93.518**	87.037	77.777
MRR	0.713	**0.722**	0.715	0.669	0.618

Table 3. Parameter w_2 study

Metric	**w_1 = 0.25, w_2, w_3 = 1.1**				
	0.15	**0.20**	**0.25**	**0.30**	**0.35**
Accuracy	92.592	93.518	92.592	87.962	82.407
MRR	0.708	0.715	0.709	0.675	0.628

Table 4. Parameter w_3 study

Metric	**w_1 = 0.25, w_2 = 0.2, w_3**		
	1.0	**1.0 + 0.10 If 1st char matches**	**1.0+ 0.20 If 1st char matches**
Accuracy	89.814	93.518	91.666
MRR	0.65	0.715	0.719

We run our model for different weightage for prefix, suffix, and similarity score and found that the best results are obtained when weightage for prefix score w1 = 0.25, suffix score w2 = 0.20, and a large value for similarity score w3 = 1.1.

The results of our model are compared with the previous model [3] that uses the Trie data structure to calculate the prefix and suffix length. Agarwal et al.'s [3] model use the similarity score directly without considering that the first character is matching or not. We had implemented the papers [3, 4] and compared their results with our proposed method. The same dataset is used to compare all the three approaches. The output of the model [3, 4] and our proposed model for a few noisy SMS tokens are shown in table 5-10. These outputs are prioritized based on the top seven retrievals obtained from the respective overall score of the three models.

ISSN: 2005-4238 IJAST
Copyright ⓒ 2020 SERSC

Table 5. Top seven words for noisy SMS word "thur"

SMS word: thur, Correct word: thursday		
Kothari's model [4]	**Agarwal's Model [3]**	**Our Proposed Model**
'thor'	'arthur'	***'thursday'***
'thud'	***'thursday'***	'tahr'
'thug'	'thud'	'thor'
'thus'	'thug'	'thou'
'tour'	'thus'	'thud'
'tahr'	'thunder'	'thug'
'thou'	'thuds'	'thus'

Table 6. Top seven words for noisy SMS word "rmv"

SMS word: rmv, Correct word: remove		
Kothari's model [4]	**Agarwal's Model [3]**	**Our Proposed Model**
'rev'	'arm'	'ram'
'ram'	'ram'	'rem'
'rem'	'rem'	***'remove'***
'rim'	***'remove'***	'rev'
'rom'	'rev'	'rim'
'rum'	'rim'	'rom'
'arm'	'rom'	'rum'

Table 7. Top seven words for noisy SMS word "dfusion"

SMS word: dfusion, Correct word: diffusion		
Kothari's Model [4]	**Agarwal's Model [3]**	**Our Proposed Model**
'fusion'	'fusion'	***'diffusion'***
'fusions'	***'diffusion'***	'infusion'
'diffusion'	'infusion'	'profusion'
'delusion'	'effusion'	'transfusion'
'effusion'	'confusion'	'effusion'
'infusion'	'suffusion'	'suffusion'
'elusion'	'perfusion'	'confusion'

Table 8. Top seven words for noisy SMS word "hxgn"

SMS word: hxgn, Correct word: hexagon		
Kothari's Model [4]	**Agarwal's Model [3]**	**Our Proposed Model**
'hag'	'sign'	***'hexagon'***
'hen'	'align'	'hexagons'
'hog'	'feign'	'hexagonal'
'hug'	'reign'	'hexane'
'hags'	'malign'	'hag'

ISSN: 2005-4238 IJAST
Copyright ⓒ 2020 SERSC

'hewn'	'assign'	'hen'
'hex'	'ensign'	'hex'

Table 9. Top seven words for noisy SMS word "sychtrst"

SMS word: sychtrst, Correct word: psychiatrist		
Kothari's Model [4]	**Agarwal's Model [3]**	**Our Proposed Model**
'schist'	'syndicalist'	***'psychiatrist'***
'scars'	'symbolist'	'schist'
'scots'	'sycophants'	'psychiatrists'
'sects'	'sycamores'	'sycophancy'
'shirt'	'sycophant'	'sycophant'
'short'	'thirst'	'sycamores'
'strut'	***'psychiatrist'***	'sycophants'

Table 10. Top seven words for noisy SMS word "decem"

SMS word: decem, Correct word: december		
Kothari's Model [4]	**Agarwal's Model [3]**	**Our Proposed Model**
'deem'	***'december'***	***'december'***
'dee'	'deem'	'deem'
'deems'	'deceit'	'deceived'
'deuce'	'decent'	'deceases'
'dace'	'decease'	'deceptions'
'deck'	'deceits'	'decelerate'
'deco'	'deceive'	'decentralised'

As we can see from the comparison table 5-8, our assumption of increasing the similarity score by 0.1 times if the first character is the same in both noisy terms and dictionary term outperforms the Agarwal et al.'s result. In table 9, we can see the better result because of decreasing the weight of the suffix score by 0.25 to 0.20. Prefix sore is also an essential factor. It works well for the truncated words. This can be seen in table 10, as for 'decem' noisy term prefix score is the most dominant factor compared to other, and as in both Agarwal's [3] method and our method, prefix score is given similar weight, i.e., 0.25. Both methods give the similar position of 'December' word.

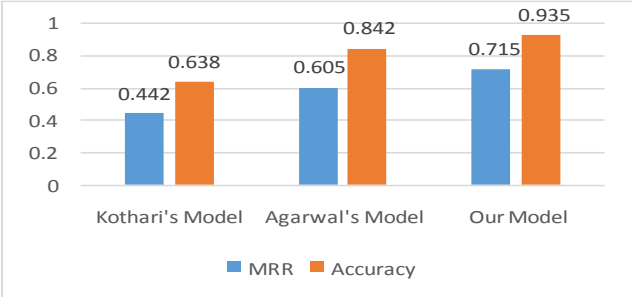


Figure 1. Mean Reciprocal Rank (MRR) and Accuracy of different models

ISSN: 2005-4238 IJAST
Copyright ⓒ 2020 SERSC

The MRR is introduced to check the effectiveness of the different models. A high value of MRR represents that the correct word corresponding to the noisy word is at the top level of the output list compare to the low MRR. Figure 1 shows the different MRR values and Accuracy of the three different models. MRR and accuracy vary with the change in the noisy data. Still, since all the three models have experimented on the same dataset, we see that our model achieves the highest value of MRR and Accuracy compared to others every time, which justifies the better performance of our proposed novel model.

5. Conclusion

Developing such an automated system to denoise an SMS text has always been a difficult task. We try to efficiently remove noise from the SMS token and achieve a good result with Accuracy being 93.518% and MRR of 0.715. We compared our results with the previous approaches and see that our model outperforms the previous models. We use the ternary tree for calculating matching prefix and suffix length as it takes less space to store the dictionary keys than the trie data structure and takes less time to search than the hash. Also, giving more weightage to the word having the first character same improves the result. As a part of future work, we can further improve the MRR and accuracy of the system and make the system work for the entire question with the best MRR score possible.

Acknowledgment

We would like to thank all the researchers for their work in SMS based FAQ retrieval. Also, we're thankful to the college for encouraging and providing support.

References

[1] P. Lalitha, "The impact of 'sms' language on standard english", IOSR Journal Of Humanities And Social Science, vol 1, 2013, pp. 50-51.

[2] L. V. Subramaniam, S. Roy, T. A. Faruquie, S. Negi, "A survey of types of text noise and techniques to handle noisy text", In Proceedings of AND, Barcelona, Spain, 2009, pp. 115–122.

[3] A. Agarwal, B. Gupta, G. Bhatt, and A. Mittal, "Construction of a Semi-Automated Model for FAQ Retrieval via Short Message Service". Proc of the 7th Forum for Information Retrieval Evaluation (FIRE2015) (2015), 35–38.

[4] G. Kothari, S. Negi, T. A. Faruquie, V. T. Chakaravarthy, and L. V. Subramaniam, "Sms based interface for faq retrieval", In the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2-Volume 2, 2009, pp. 852-860.

[5] E. Sneiders, "Automated question answering using question templates that cover the conceptual model of the database", Natural Language Processing and Information Systems, Springer Berlin Heidelberg, 2002, pp. 235-239.

[6] D. Shaikh, M. Jain, M. Rawat, R. R. Shah, and M. Kumar, "Improving accuracy of sms based faq retrieval system", In Multilingual Information Access in South Asian Languages, Springer, 2013, pp. 142-156.

[7] D. Contractor, T. A. Faruquie, and L. V. Subramaniam, "Unsupervised cleansing of noisy text", In Proceedings of the 23rd International Conference on Computational Linguistics: Posters, 2010, August, pp. 189-196.

[8] N. Shivhre, "Sms based faq retrieval", In Multilingual Information Access in South Asian Languages, Springer Berlin Heidelberg, 2013, pp. 131-141.

[9] S. Mhaisale, S. Patil, and K. Mahamuni, "Weighted edit distance based faq retrieval using noisy queries", Post-Proceedings of the 4th and 5th Workshops of the Forum for Information Retrieval Evaluation (FIRE2012&2013) (2013), 8:1–8:4.

[10] R. Batra, S. Sharma, A. Shrivastav, P. Goyal, "Efficiently denoising sms text for faq retrieval", IEEE International Conference on Data Mining and Intelligent Computing, Delhi, Sep. 2014.

[11] CornCorb. The Corncob list of more than 58000 English words. Available: http://www.mieliestronk.com/wordlist.htm

ISSN: 2005-4238 IJAST
Copyright ⓒ 2020 SERSC

View publication stats

J Mater Sci: Mater Electron

Rapid detection of biomolecules in a dielectric modulated GaN MOSHEMT

Shaveta[1,2], H. M. Maali Ahmed[1], and Rishu Chaujar[1,*]

[1] *Department of Applied Physics, Delhi Technological University, Main Bawana Road, Delhi, India*
[2] *Solid State Physics Laboratory, Defence Research Development Organization, Delhi, India*

Received: 2 June 2020
Accepted: 8 August 2020

© Springer Science+Business Media, LLC, part of Springer Nature 2020

ABSTRACT

Biosensors are the devices that find application in almost every field nowadays. In this paper, GaN MOSHEMT based biosensor is proposed for detection of biomolecules such as ChOx, protein, streptavidin and Uricase. The effect of biomolecule species on the performance parameters of the device has been studied. It has been observed that there is a significant increase in the drain current and g_d is observed with the addition of biomolecule in the nanocavity. The electron concentration contour is studied which shows the rise of carrier concentration with biomolecule. Maximum positive shift is observed in threshold voltage for Uricase due to lowest dielectric constant. Similarly, the change in transconductance is also obtained with biomolecules. The effect of cavity dimensions on sensitivity is also studied. The maximum increase of 10% in channel potential is noted due biomolecule presence in the cavity. This device has shown good sensing and can be used for biosensing applications efficiently in addition to the high power performance of MOS-HEMTs.

1 Introdcution

The recent COVID-19 pandemic has again brought in forefront the importance of electronic biosensors in detection of biomolecules & biological warfare agents. The electronic biosensors may find applications in real time monitoring of airborne biomolecules in transport systems like aircrafts & metro rail and centrally air conditioned buildings like schools, hospitals, etc. Therefore, the need of the hour is to design and develop fast and accurate biosensors.

Biosensors use chemical reactions to detect the biochemical compounds such as antibodies, biological molecules, nucleic acids, enzymes etc. So they are widely used in many applications such as monitoring of diseases, food analysis, crime detection, environmental field monitoring, and also for the study of biomolecules interaction [1, 2].

Now a days Gallium Nitride (GaN) is emerging as one of the most promising material due to very promising properties. GaN has wide band gap of 3.4 eV and can operate at high temperature (> 300 °C) as compare to silicon, which is being used to fabricate conventional devices. Apart from this they have high electric field, polarization nature, distinctive chemical stability in the biological

Address correspondence to E-mail: chaujar.rishu@dtu.ac.in

https://doi.org/10.1007/s10854-020-04216-7
Published online: 31 August 2020

atmosphere, low level of toxicity to the living cells, and the ability to integration with other electronic devices [3–7]. AlGaN/GaN HEMT shows promising biological/chemical sensing applications due to the presence of 2DEG (Two-dimensional electron gas) at the hetero interface. Any biomolecules can get easily attached to AlGaN barrier layer that varies the surface charges at the AlGaN/GaN interface and as a counter effect channel properties are varied. So interface property [8] is very crucial parameter in determining the device functionality. So HEMT based gas sensors, mercury and chlorine ion sensors, pH sensors, DNA sensors, glucose sensor, cancer biomarker detector, terahertz sensors, strain sensors etc. have been developed [9, 10].

Extensive work has been done by various researchers on FET based biosensors for detection of various biomolecules. Schwarz et al. have reported a novel HEMT-based hybridization sensor using photochemical functionalization to detect DNA [11]. Khazanskaya et al. have reported field-effect transistor biosensor for ammonium ions and urea detection [12]. Chen et al. used Au-gated AlGaN/GaN HEMTs to detect c-erbB-2, which is a breast cancer marker [13]. Kannan reported GaN HEMT to detect pH and biomarkers for the detection of biomolecules using MOS transistor sensor [14]. Yantao et al. have invented an effective FET biosensor for Ebola Antigen detection [15]. Chung et al. demonstrated AlGaN/GaN HEMT hydrogen gas sensor at extreme environmental conditions [16]. Hu proposed AlGaN/GaN/AlN quantum-well MOS-HEMTs based on numerical simulations [17]. Recently, Seo et al. reported a field-effect transistor (FET)-based biosensing device by using graphene sheet for detecting SARS-CoV-2 in clinical samples [18].

In this paper, Simulation of GaN MOSHEMT based biosensor for detection of neutral biomolecule such as ChOx, protein, streptavidin and Uricase has been demonstrated. Although gate-all-around junctionless transistor (GAA-JLT) [19] and inversion mode MOSFETs [20] have been demonstrated in past for label-free electrochemical detection of neutral biomolecule. As compare to these devices GaN MOSHEMT devices have advantages of reduced gate leakage current density, high breakdown voltage and high electron mobility under both low and high transverse fields [21, 22]. The structure is analyzed by using dielectric modulation technique. Silvaco TCAD software is used for simulation of the structure.

Section 2 presents the device architecture and models used for simulation. Section 3 discusses the device current model used for simulation. Results including output and transfer characteristics, sensitivity and channel potential are discussed in Section 5.

2 Device architecture

Figure 1a highlights the schematic cross-section of AlGaN/GaN MOSHEMT. The structure has 1 μm thick GaN buffer layer, 1 nm AlN spacer, 20 nm oxide, 18 nm $Al_{0.20}Ga_{0.80}$ N barrier. The gate length (L_G) is 5 μm, length between source and gate (L_{SG}) is 1 μm and between gate and drain (L_{GD}) is 2 μm. The structure is passivated by thin film of Si_3N_4 layer to reduce surface states effect. The complete length and width of the device is 9 μm and 100 μm, respectively. Contact taken at the gate terminal is Schottky and at source & drain is ohmic. Initially the region below the cavity was filled with SiO_2 and later two cavities of 1.5 μm each are considered in region I and III underneath the gate electrode as shown in Fig. 1b. This region is confined on the top of AlGaN layer to induce and detect biomolecule. In this paper, neutral biomolecules like ChOx, Protein, Streptavidin and Uricase are introduced in the cavity. The dielectric constant of these biomolecules is given in Table 1.

Considered neutral biomolecules do not react with the SiO_2 interface. They only introduce the dielectric in the region and the structure is like gate-stack.

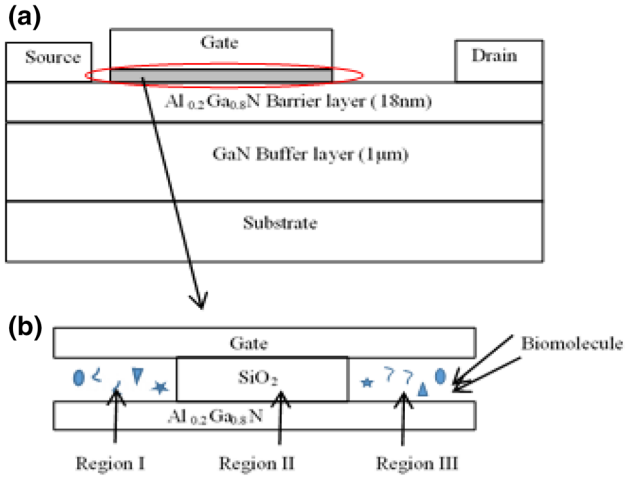


Fig. 1 **a** Schematic of AlGaN/GaN MOSHEMT with nanocavity for biomolecules. **b** Magnified view of cavity region introduced below gate region

Therefore, the presence of neutral biomolecules can be modeled by introducing material with a dielectric constant corresponding to neutral biomolecules. This approach has been calibrated with experimental results [23].

All simulations are carried out in Silvaco TCAD software. The models used are polarization model, Shockley–Read–Hall (SRH) recombination mode, field dependent drift velocity (FLDMOB) model, concentration dependent mobility (CONMOB), albrct.n and 8.85 × 10^{12} cm^{-2} of interface charges are considered at the AlGaN/GaN interface.

3 Device current model

As shown in Fig. 1b region I and III is empty cavity to be filled with biomolecule and region II is oxide layer. The capacitance for different regions is given by Eqs. (1) and (2)

$$C_{SiO_2} = \frac{\varepsilon_{SiO_2}}{t_{ox}} \tag{1}$$

$$C_{Bio} = \frac{2\varepsilon_{Bio}}{t_{Bio}} \tag{2}$$

where, ε_{SiO2}, ε_{Bio}, t_{ox} and t_{Bio} are permittivity and thickness of SiO_2 and biomolecule, respectively. The thickness of oxide and biomolecule layer is kept constant for simplicity.

The total capacitance of the region I, II and III is given by Eq. (3)

$$C_{Total} = C_{SiO_2} + C_{Bio} \tag{3}$$

Total capacitance of the MOSHEMT device is given by Eq. (4)

$$\frac{1}{C_{MOSHEMT}} = \frac{1}{C_{HEMT}} + \frac{1}{C_{total}} \tag{4}$$

where, C_{HEMT} is the capacitance of the Device [24]. The drain current for the device is given by Eq. (5) Ref.[25],

Table 1 Dielectric constant of biomolecules

Biomolecule	Permittivity
ChOx	3.5
Protein	2.5
Streptavidin	2.1
Uricase	1.5

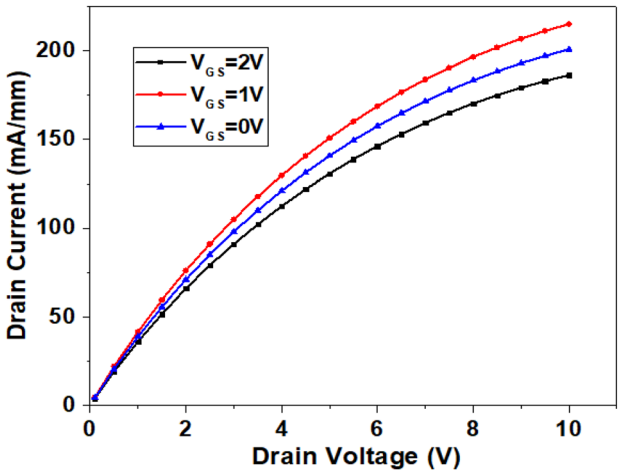


Fig. 2 Output characteristics of AlGaN/GaN MOSHEMT without biomolecules at different gate voltages

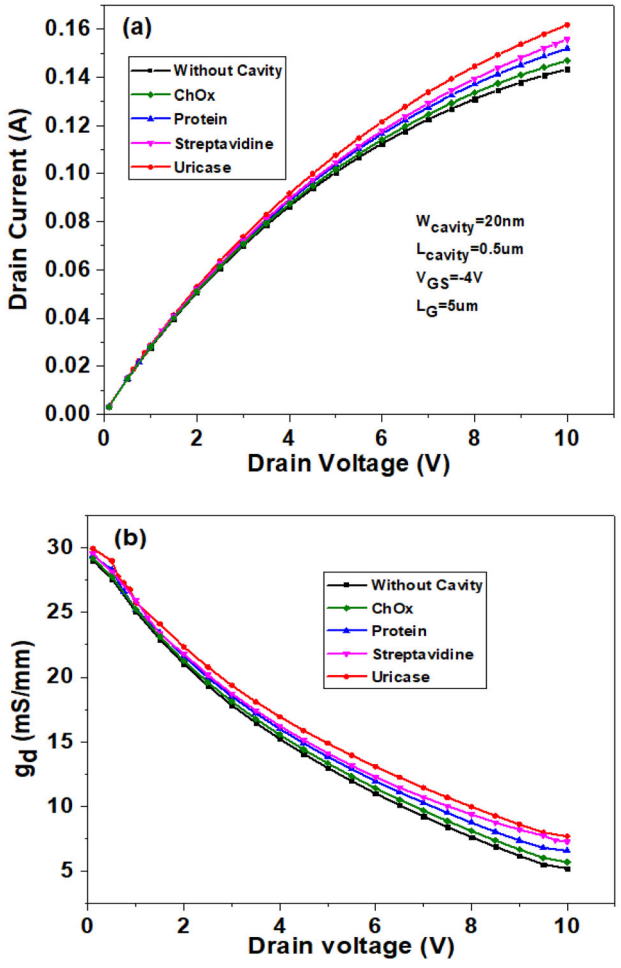


Fig. 3 **a** Output characteristics **b** g_d vs Vd of AlGaN/GaN MOSHEMT with different biomolecules introduced in the cavity

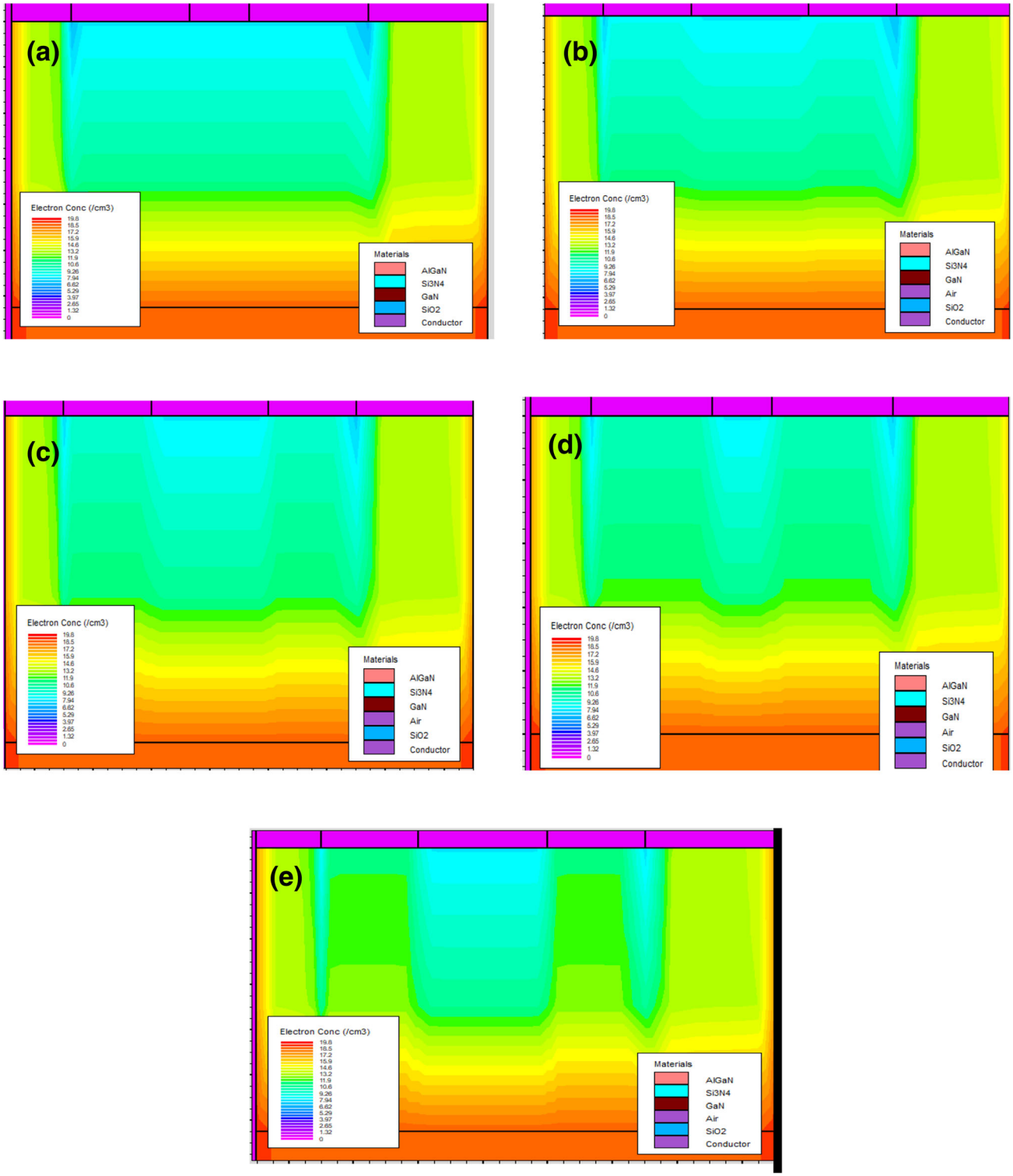


Fig. 4 Electron concentration contour (**a**) without biomolecule (**b**) ChOx (**c**) Protein (**d**) Streptavidin (**e**) Uricase

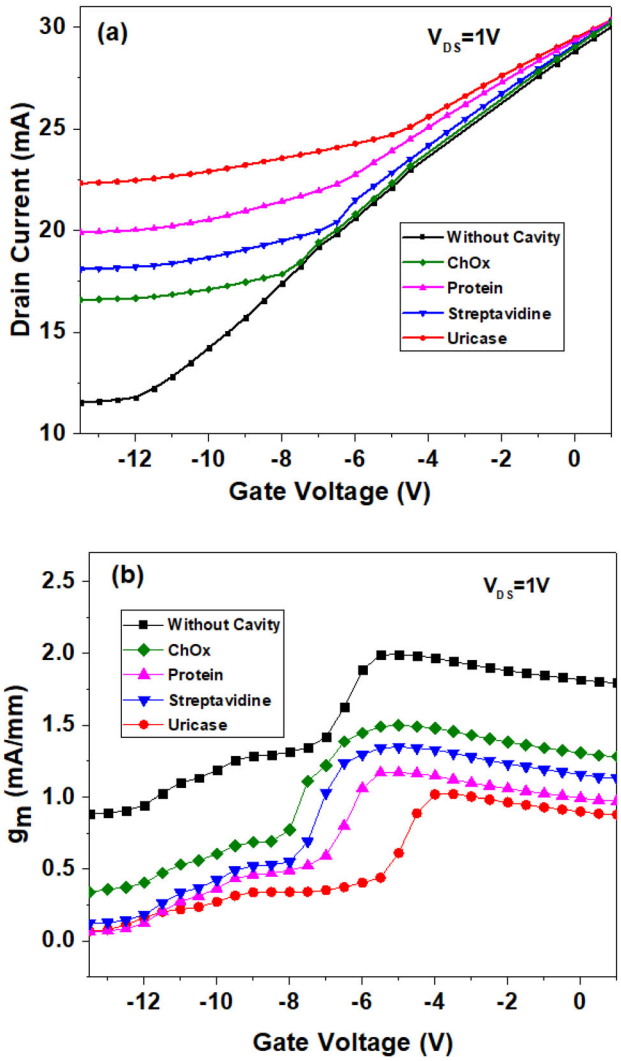


Fig. 5 **a** Output characteristics **b** g_d vs Vd of AlGaN/GaN MOSHEMT with different biomolecules introduced in the cavity

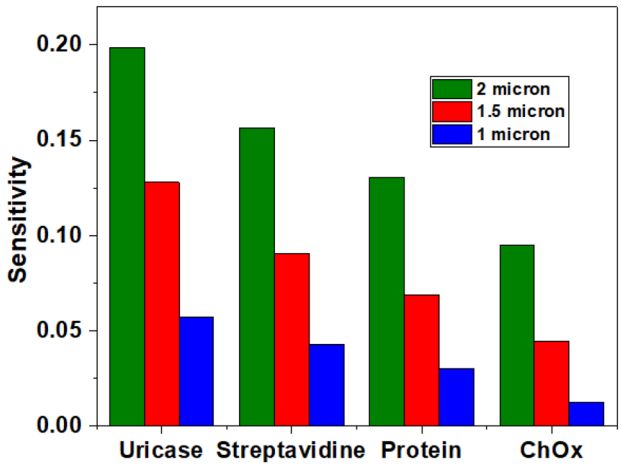


Fig. 6 sensitivity parameter for different biomolecules

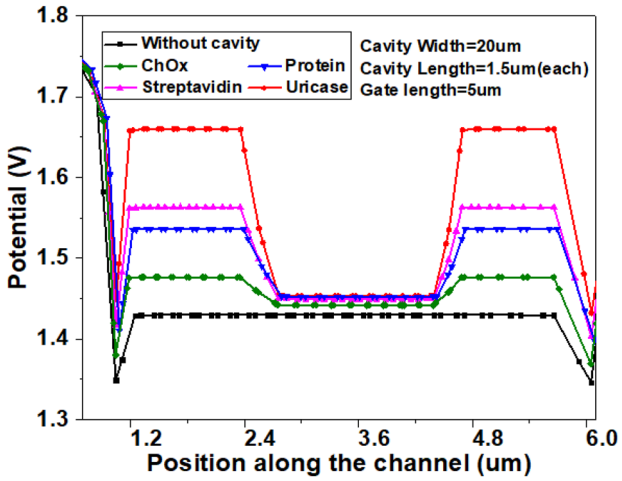


Fig. 7 Channel potential for biomolecules along the channel length

$$I_{D=-q\mu\frac{W_G}{L_G}}\left(\begin{array}{c}\frac{2}{s}\gamma_0\left(n_{drain}^{\frac{5}{s}}-n_{source}^{\frac{5}{s}}\right)+v_{th}\left(n_{drain}-n_{source}\right)\\+\frac{q}{2C_{MOSHEMT}}\left(n_{drain}^2-n_{source}^2\right)\end{array}\right) \quad (5)$$

$\gamma_o = 2.12 \times 10^{-12}$ V $m^{4/3}$ is constant. n_{source} and n_{drain} are charge concentration at source and drain, respectively. W_G is width and L_G is gate length. μ is the low field mobility.

4 Results and discussions

Figure 2 shows the Id-Vd characteristics of the GaN based MOSHEMT having gate length of 5 μm without biomolecule. The core current models used in the simulation are able to reproduce the device characteristics very well. The self-heating effects are not considered in the simulation.

The biomolecules has been introduced in cavities of Wcavity = 20 μm and Lcavity = 1.5 μm on both sides of oxide layer. Figure 3a shows the output characteristics of the device. When biomolecule is introduced in to the cavity, there is enhancement in the carrier concentration in the channel region which causes increase in the drain current. Here Maximum increase of 18 mA in drain current is observed for Uricase. g_d vs Vd plot is shown in Fig. 3b. Since Id-Vd characteristics of the device are changing by presence of biomolecule in the cavity in similar way variation in g_d is also observed.

Table 2 Comparison between MOSHEMT and GAA-JLT [26]

Biomolecule	ChOx	Protein	Streptavidine	Uricase	
Change in current (mA/mm) = current without cavity − current with biomolecule	MOSHEMT (ΔI)	136	224	280	350
	GAA-JLT (ΔI_{off})	0.0033	0.029	0.0538	0.170
Change in threshold voltage (ΔV_{th})	MOSHEMT	4.027	6.614	8.360	10.844
	GAA-JLT	0.038	0.093	0.118	0.150

Figure 4a–e shows the electron concentration contour. Figure 4a shows the no change of concentration when no biomolecule is introduced. As biomolecule is introduced, the more charges gets induced in the channel in the region beneath the cavity. The quantity of charge induced depends on the dielectric constant of the biomolecule as can be seen from Fig. 4b–e. Due to lowest value of dielectric constant in Uricase more charges are induced in the channel that is inferred from the rise of drain current in Fig. 3a.

Transfer characteristics and transconductance are shown in Fig. 5a, b. The positive shift in the threshold voltage is observed when biomolecule is introduced in the cavities. Because the presence of biomolecule induces charges in the channel thus results in the better formation of 2DEG in the channel at lower gate voltage. The Threshold voltage shifted from − 12 V for without biomolecule to − 5.5 V at V_{DS} = 1 V for Uricase. However, with increase of dielectric constant for streptavidin, protein and ChOx negative shift in the voltage is observed as compared to Uricase the due to reduced induced charges. Similarly the change in the transconductance is also observed as shown in Fig. 4b. The maximum transconductance is 2 mA/mm for without biomolecule, 1.5 mA/mm for ChOx, 1.3 mA/mm for Protein, 1.2 mA/mm for streptavidin and 1 mA/mm for Uricase.

Figure 6 shows the sensitivity parameter for different biomolecules. The maximum increase of 0.141 in sensitivity is observed when cavity length is varied from 1 to 2 µm on both each side of the oxide layer. This is due to reason, as cavity length increases, more area of the biomolecule is available to interact with underneath AlGaN functional layer. The maximum increase is observed for Uricase due to its lower dielectric constant. So for high sensitivity applications, the length of the cavity should be more.

Figure 7 shows the potential along the channel at 4 µm below the AlGaN/GaN interface when biomolecule is introduced in the cavity region I&III underneath the gate. Permittivity of the biomolecules is different due to which the potential developed in the channel region changes. Uricase having lowest permittivity is showing potential variation of 10% as compared to 3% for ChOx. The comparison in parameter like on current and threshold voltage for MOSHEMT and GAA-JLT [26] for detection of neutral biomolecule is shown in Table 2. A large change in drain current and threshold voltage of MOSHEMT is observed for all biomolecules.

5 Conclusion

In this paper, the used model accurately predicts the effect of biomolecule species on the performance parameters of GaN based MOSHEMT. There is significant increase in the drain current and change g_d is observed with the addition of biomolecule in the cavity. The electron concentration contour indicates notable increase of electron concentration with addition of biomolecule. The positive shift in the gate threshold voltage, maximum shift of ∼ 54% is observed for Uricase due to lowest dielectric constant. Similarly the notable change in transconductance is also obtained. The sensitivity of the structure increases approximately four times with the increase in the cavity length due to more area of contact of the biomolecule with the underneath functional layer. The rise of 10% in channel potential is viewed when biomolecule (Uricase) is introduced in the cavity. MOSHEMT showed better sensitivity towards drain current and threshold voltage as compared to GAA-JLT. Therefore, it can be concluded that this GaN Based MOSHEMT gives significant change in performance parameters and can be used for biomolecule sensing applications effectively.

Acknowledgements

We would like to thank our guide for motivation, constant guidance and all technical support. Our sincere regards to faculty of Applied Physics

Department, Delhi Technological University, New Delhi.

References

1. J.-Y. Yoon, *Introduction to Biosensors: From Electric Circuits to Immunosensors* (Springer, New York, 2016)
2. A. Kumar, M. Roy, N. Gupta, M.M. Tripathi, R. Chaujar, Dielectric modulated transparent gate thin film transistor for biosensing applications. Mater. Today (2020).
3. F. Ren, J.C. Zolper, *Wide Band Gap Electronic Devices* (World Scientific publication Co. Pte. Ltd, Singapore, 2003)
4. U.K. Mishar, P. Parikh, Y. Wu, AlGaN/GaN HEMTs—an overview of device operation and applications. Proc. IEEE **90**(6), 1022–1031 (2002)
5. O. Ambacher, J. Smart, J. Shealy, N. Weimann, K. Chu, M. Murphy, W. Chaff, L. Eastman, R. Dimitrov, L. Wittmer, M. Stutzmann, W. Rieger, J. Hilsenbeck, Two-dimensional electron gases induced by spontaneous and piezoelectric polarization charges in N- and Ga-face AIGaN/GaN heterostructures. J. Appl. Phys. **85**, 3222–3233 (1999)
6. J. Madan, H. Arora, R. Pandey, R. Chaujar, Analysis of varied dielectrics as surface passivation on AlGaN/GaN HEMT for analog applications. in *International Conference on Wireless Networks & Embedded Systems (WECON-2018)*, p. 15–18, 16–17
7. H. Arora, J. Madan, R. Chaujar, Impact on analog and linearity performance of nanoscale AlGaN/GaN HEMT with variation in surface passivation stack. Mater. Today (2018)
8. T. Aggerstam, S. Lourdudoss, H.H. Radamson, M. Sjödin, Investigation of the interface properties of MOVPE grown AlGaN/GaN high electron mobility transistor (HEMT) structures on sapphire. Thin Solid Films **515**, 705 (2006)
9. V.K. Khanna, Robust HEMT microsensors as pospective successors of MOSFET/ISFET detectots in harsh environment. Front. Sens. **1**(3), 38–48 (2013)
10. M. Stutzmann et al., GaN-based heterostructures for sensor applications. Diamond Relat. Mater. **11**(3), 886–891 (2002)
11. S.U. Schwarz et al., DNA-sensor based on AlGaN/GaN high electron mobility transistor. Phys. Status Solidi A **208**(7), 1626–1629 (2011)
12. N. Kazanskaya et al., FET-based sensors with robust photosensitive polymer membranes for detection of ammonium ions and urea. Biosens. Bioelectron. **11**(3), 253–261 (1996)
13. K.H. Chen, B.S. Kang, H.T. Wang, T.P. Lele, F. Ren, D.M. Dennis, K.J. Linthicum, C-erbB-2 sensing using Al Ga N/Ga N high electron mobility transistors for breast cancer detection. Appl. Phys. Lett. **92**, 192103 (2008). https://doi.org/10.1063/1.2926656
14. N. Kannan, M.J. Kumar, Charge-modulated underlap I-MOS transistor as a label-free biosensor: a simulation study. IEEE Trans. Electron Devices **62**(8), 2645–2651 (2015)
15. Y. Chen et al., Field-effect transistor biosensor for rapid detection of Ebola antigen. Sci. Rep. **7**(1), 1–8 (2017)
16. G.H. Chung, T.A. Vuong, H. Kim, Demonstration of hydrogen sensing operation of AlGaN/GaN HEMT gas sensors in extreme environment. Physics **12**, 83–84 (2019). https://doi.org/10.1016/j.rinp.2018.11.064
17. W.D. Hu, X.S. Chen, Z.J. Quan, X.M. Zhang, Y. Huang, C.S. Xia, W. Lu, P.D. Ye, Simulation and optimization of GaN-based metal-oxide-semiconductor high-electronmobility-transistor using field-dependent drift velocity model. J. Appl. Phys. **102**, 034502 (2007). https://doi.org/10.1063/1.2764206
18. G. Seo, G. Lee, J.-O. Lee, B.T. Kim, E.C. Park, S.I. Kim, Rapid detection of COVID-19 causative virus (SARS-CoV-2) in human nasopharyngeal swab specimens using field-effect transistor based biosensor. Am. Chem. Soc. **14**(4), 5135–5142 (2020). https://doi.org/10.1021/acsnano.0c02823
19. T. Wang, L. Lou, C. Lee, A junctionless gate-all-around silicon nanowire FET of high linearity and its potential applications. IEEE Electron Devices Lett. **34**(4), 478–480 (2013)
20. L. Xiang, Z. Chen, N. Shen, N. Singh, K. Banerjee, G.Q. Lo, D.L. Knong, Vertically stacked and independently controlled twin-gate MOSFETs on a single silicon nanowire. IEEE Electron Devices Lett. **32**(11), 1492–1494 (2011)
21. S. Taking, D. MacFarlane, E. Wasige, AlN/GaN-based MOS-HEMT technology: processing and device results. Act. Passiv. Electron. Compon. (2011). https://doi.org/10.1155/2011/821305
22. K. Han, L. Zhu, GaN MOSHEMT employing HfO2 as a gate dielectric with partially etched barrier. Semicond. Sci. Technol. **32**(9), 095004 (2017)
23. I.M. Hyungsoon, X.-J. Huang, B. Gu, Y.-K. Choi, A dielectricmodulated field-effect transistor for biosensing. Nat. Nanotechnol. **2**, 430–434 (2007)
24. Y. Yuan-Zheng, H. Yue, F. Qian, N. Jin-Yu, GaN MOS HEMT using Ultra thin Al_2O_3 dielectric grown by atomic layer deposition. Chin. Phys. Lett. **24**(8), 2419 (2007)
25. F.M. Yigletu, S. Khandelwal, T.A. Fjeldly, B. Iñiguez, Compact charge-based physical models for current and capacitances in AlGaN/GaN HEMTs. IEEE Trans. Electron Devices **60**(11), 3746–3752 (2013)
26. Y. Pratap, M. Kumar, S. Kabra, S. Haldar, R.S. Gupta, M. Gupta, Analytical modeling of gate-all-around junctionless transistor based biosensors for detection of neutral biomolecule species. J. Comput. Electron. **17**, 288–296 (2018)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of *Functional Biomaterials*

MDPI

Review

Recent Advances of Conducting Polymers and Their Composites for Electrochemical Biosensing Applications

John H.T. Luong [1,*], Tarun Narayan [2], Shipra Solanki [2,3] and Bansi D. Malhotra [2]

[1] School of Chemistry and the Analytical and Biological Chemistry Research Facility (ABCRF), University College Cork, College Road, Cork T12 YN60, Ireland
[2] Department of Biotechnology, Delhi Technological University, Delhi 110042, India; narayantarun41@gmail.com (T.N.); shiprasolanki.y29@gmail.com (S.S.); bansi.malhotra@gmail.com (B.D.M.)
[3] Applied Chemistry Department, Delhi Technological University, Delhi 110042, India
* Correspondence: luongprof@gmail.com

Received: 31 August 2020; Accepted: 20 September 2020; Published: 25 September 2020

Abstract: Conducting polymers (CPs) have been at the center of research owing to their metal-like electrochemical properties and polymer-like dispersion nature. CPs and their composites serve as ideal functional materials for diversified biomedical applications like drug delivery, tissue engineering, and diagnostics. There have also been numerous biosensing platforms based on polyaniline (PANI), polypyrrole (PPY), polythiophene (PTP), and their composites. Based on their unique properties and extensive use in biosensing matrices, updated information on novel CPs and their role is appealing. This review focuses on the properties and performance of biosensing matrices based on CPs reported in the last three years. The salient features of CPs like PANI, PPY, PTP, and their composites with nanoparticles, carbon materials, etc. are outlined along with respective examples. A description of mediator conjugated biosensor designs and enzymeless CPs based glucose sensing has also been included. The future research trends with required improvements to improve the analytical performance of CP-biosensing devices have also been addressed.

Keywords: biosensor; conducting polymer; enzymeless; electropolymerization; polypyrrole; polyaniline; nanomaterials; glucose

1. Introduction

Electroanalysis is commonly performed on a flat electrode of noble metals (gold, platinum, palladium, etc.), carbon-based materials (glassy carbon, highly ordered pyrolytic graphite, boron-doped diamond, etc.), indium tin oxide (ITO), screen printed electrode (SPE), etc. For enhanced sensitivity, the electrode surface is modified with metal nanoparticles, carbon nanotubes, carbon monoliths, quantum dots, graphene/its derivatives, and conducting polymers (CPs). A review of the use of advanced nanomaterials in biosensor construction is available elsewhere [1]. The selective signal response is based on a specific biorecognition element, which is confined to the sensing area of the electrode by physical adsorption or covalent bonding. In this context, CPs with excellent electrical conductivity, high electronic affinity, and low ionization potential play several key roles in biosensor performance. They are easily polymerized on the electrode surface by a noble salt or a dopant to augment conductivity, catalytic activity, and high surface area. High conductivity is always a desired feature to minimize the interface resistance between the electrolytes, resulting in high signal

response. In a broad sense, CPs encompass π-conjugated polymers and polymers whose backbone has alternating single and double (or triple) covalent bonds. Since their discovery, CPs have been extensively used in the biomedical field for biosensor fabrication and tissue engineering. The non-conducting polymers can be used to give mechanical support to the sensor membranes, but CPs improve the sensitivity and stability of the sensing devices along with providing strength. Another property that makes CPs ideal matrices for biosensors is their excellent biocompatibility; CPs have shown extremely stable immobilization of biomolecules on their surface with complete retention of their activity. The most common methods used are cross-linking, covalent binding, physical entrapment, and adsorption [2].

Polymers with amino groups (-NH_2) and carboxyl groups (-COOH) are easily bioconjugated with the required biorecognition molecules such as enzymes, antibodies, and DNA. Thiolated biomolecules, mostly DNA, self-assemble on conducting polymers decorated with gold or silver nanoparticles. Conducting polymers can be incorporated with hydrogels swollen with water to form conducting porous polymer hydrogels with several desired features for biosensing platforms [3]. Conducting porous polymer hydrogels exhibit ultra-high conductivity, high mechanical properties, hierarchical structure, and swelling nature. High porosity facilitates the diffusion of the analyte and its product(s) in and out of the sensing electrode. Hydrogels swell with water, are biocompatible and serve as an excellent milieu to accommodate and prolong the activity of fragile recognition molecules.

Among four classical conducting polymers, polypyrrole (PPY) and its derivatives have been used extensively in biosensor construction. They are biocompatible and protect electrodes from fouling and rarely cause any significant disturbance of the working environment. In some cases, they form perm-selective films to exclude endogenous electrochemically active interferents. Polyaniline (PANI), another common conducting polymer, loses its conductivity at pH levels above 4, so it is commonly used with another conducting polymer, nanoparticles, or carbon nanomaterials, mainly carbon nanotubes (CNT) and graphene.

Electrochemical polymerization (EP) has three different steps: (i) solubilization of oligomers in the diffusion layer after monomer oxidation, (ii) their deposition by nucleation followed by growth, and (iii) chain propagation by solid-state polymerization [4]. However, a general mechanism for each polymer cannot be easily established as the process is governed by operating parameters, except for the initial oxidation step [4,5]. The properties of electropolymerized films are dependent upon electrode materials, dopants, electrolyte, monomers, pH, and electrochemical deposition methods. Comprehension of nucleation and growth kinetics allows for the tailoring of polymer properties like density, morphology, and crystallinity, according to requirement. CPs with carboxyl or amino groups serve as immobilization matrices for covalent attachment of recognition molecules [6–9]. Some polymers like polyphenanthroline facilitate the direct electron transfer (DET) from the active site of biomolecules and the electrode [10]. In some cases, a composite of two polymers, PPY and PANI, are used together to anchor the enzyme on the electrode [11].

This review paper provides an updated overview of CPs based biosensing platforms, based on publications during 2018 to 2020. Literature searches frequently encounter several claims for the choice of electrochemical biosensors and immunosensors: simplicity, high sensitivity, robustness, mass production, miniaturization, multiplexing, and portability. A critical assessment of their analytical performances will be addressed, whether CPs can provide an attractive paradigm in terms of preparations and applications. Except for the non-enzymatic detection of glucose, the use of CPs without a biorecognition molecule to form a chemosensor is excluded in this review regardless of whether the target analyte is a biomolecule. To facilitate reading, biosensing platforms based on conducting polymers, polymer composites, and direct electron transfer are discussed separately. The analytical performances of such biosensors are also tabulated in three different tables for comparison.

2. Electrochemical Polymerization

Conducting monomers can be polymerized by chemical or electrochemical methods. The former produces powders in bulk and is suitable for mass production while the latter results in thin films on the electrode surface with controlled thickness and characteristics. In EP, a solvent consisting of the desired monomer and an anionic doping salt is subjected to an anodic potential. The monomer is oxidized to form the radical cation, which reacts with other neighboring monomers leading to the formation of oligomeric products and eventually polymer deposition at the anode. Both the selected solvent and the doping salt must be stable at the oxidation potential required for EP. The choice of the solvent depends upon the target monomer, e.g., pyrrole is water-soluble, but thiophene is not. Acetonitrile and propylene carbonate are commonly used as the solvent due to their high relative permittivities and permit the use of very large potential windows. EP is commonly performed by cyclic voltammetry (CV), potentiostatic/galvanostatic polarization. The cyclic potential sweep procedure is simple but time-consuming due to dead time in potential regions where no polymerization occurs.

Polypyrrole (PPY): *Synthesis and Mechanisms*. Pyrrole (PY, MW 67.09, C_4H_5N, $pK_a = 4$) is water-soluble (45 mg/mL, 0.67 M at 25 °C) [12]and colorless when fresh. It is also soluble in alcohol, ether, dilute acids, and in most organic chemicals [13]. Its corresponding polymer (PPY) can be easily prepared by chemical oxidative polymerization. Thus, the simple casting of PPY by evaporation of a PPY solution is simple, but the resulting film is not compact or uniform. It is also challenging to control the exact amount of PPY on the active electrode area and the film thickness. Thus, simple casting is excluded for the preparation of isolated microelectrodes and their array. In contrast, EP creates a uniform film on only electroactive surfaces with high reproducibility regardless of their dimensions. PPY is widely used in biosensor construction as the film offers several desired features and PY is commercially available. The classical CV technique offers qualitative information on the redox processes in the polymerization initially and is useful to characterize the polymer films. Potentiostatic techniques can be conducted at a constant potential or a constant current mode to provide more quantitative information to elucidate the nucleation mechanism and macroscopic growth.

Among several suggested mechanisms for the formation of PPY, the mechanism proposed by Diaz et al. [14] is frequently encountered in the literature. This mechanism deciphers several successive reactions: radical cation formation, radical coupling, and deprotonation (Scheme 1). It is challenging to probe the propagation and termination steps during EP as intermediate high oligomers are insoluble. The film thickness and morphology deserve a brief discussion here because these important parameters control the diffusion of the target analyte and its corresponding product(s) in and out of the film layer. A supporting electrolyte containing potassium iodide (KI) yields a relatively thick (~0.6 μm) and highly conductive film on a glassy carbon electrode (GCE), whereas KF produces a very thin film (<0.0001 μm) with negligible conductivity [15]. The film thickness is ~0.03–0.04 μm when either $KClO_4$, KCl, KBr, or KNO_3 is included in the electrolyte, compared to <0.01 μm with K_2SO_4. EP is performed by CV with 40 cycles from −0.3 to 0.9 V vs. Ag/AgCl, at 0.1 V/s in 0.1 M PY with 0.1 M of the above salts. The film thickness only increases very slowly after the second cycle of CV. The film thickness is also governed by different electrochemical methods and their associated operating parameters. As an example, the PPY film thickness obtained from electrosynthesis at a constant potential of 0.90 V vs. Ag/AgCl, at 20 s and 300 s is about ~2 μm and >25 μm in the electrolyte consisting of 0.1 M PY and 0.1 M potassium iodide [15]. Except for the use of KF, other salts at a concentration of 0.1 M provide a film with thickness ranging from above 5 μm to ~8 μm. In situ atomic force microscopy (AFM) is a great tool to probe the morphology and thickness of the film. For thicknesses below 1 μm, the film exhibits a globular-shaped structure and is independent of the dopant used. Thicker chloride-doped and perchlorate-doped films show cauliflower's structures [16].

PPY films prepared at lower temperatures have poorer adhesion with an uneven appearance as compared to those prepared at higher temperatures as highly conductive PPY films (500 S/cm) can be prepared at 10 °C [17]. However, if the recognition biomolecule is entrapped by the PPY film in a

one-step reaction, the experiment must be conducted at a moderate temperature. The PPY stability in aqueous solution depends on pH as it undergoes protonation in strong acids (pKa of 2–4) and deprotonation in a basic solution (pK_a of 9–11). The conductivity decreases from 10 S/cm to 0.1 S/cm when the pH increases from below 7 to pH 11. In most biosensor applications, the recognition biomolecules are compatible at physiological pH. The effect of water on the properties of PPY is another important point. Water exhibits stronger basicity than PY and captures the H^+ generated during the EP to prevent the formation of the PY trimer [18]. Solvents like methanol, ethanol, and tetrahydrofuran, are more basic as compared to PY and are not capable of deprotonating the intermediate radical cations. Thus, their use is as effective as water to avoid the passivation of the electrode by preventing the plausible formation of the PY trimer. Additionally noteworthy, is the over oxidation of PY, which results in a non-conductive film. For repeated analyses, the PPY film must be firmly attached to the electrode surface, therefore, the deposition of a conducting polymer must be carefully performed and optimized accordingly.

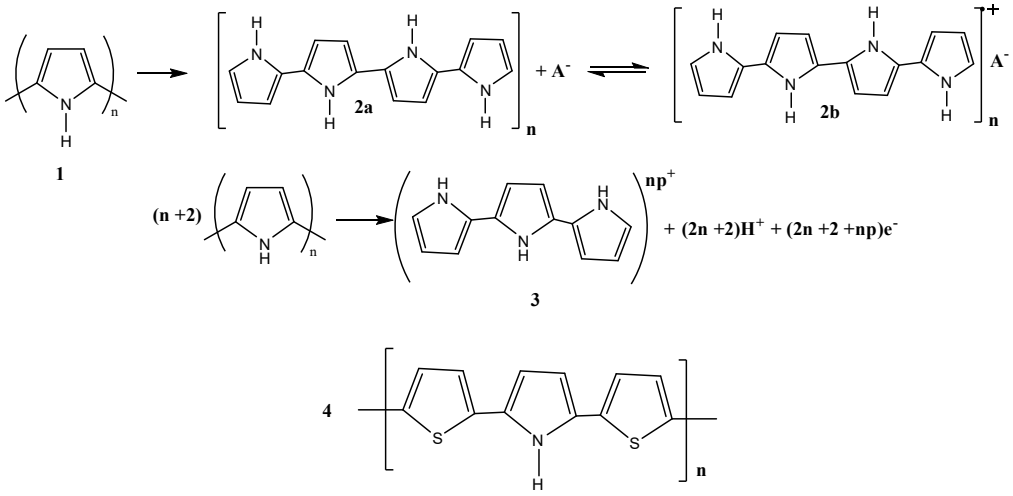
Among several commercial PY derivatives, 4-(3-pyrrolyl)butyric acid and 3-(1*H*-pyrrol-1-yl)-1-propanamine are useful for bioconjugation with the NH_2 or COOH groups of recognition biomolecules via -NH_2 bonding, glutaraldehyde crosslinking or carbodiimide covalent bonding between the -NH_2 and -COOH groups [19]. An electropolymerized pyrrolepropylic acid (PPA) film with a high density of -COOH groups is used to covalently attached protein probes, leading to significantly improved detection sensitivity compared with conventional entrapment methods [9]. Poly(*N*-methyl pyrrole), synthesized from *N*-methyl pyrrole, exhibits a conductivity three orders of magnitude lower than that of PPY [14]. Poly(2,5-di-(-2-thienyl)-pyrrole can be doped with para-toluene sulfonate with electrical conductivities between 10^{-8} and 10^{-1} S/cm [20].

Polyaniline (PANI): Synthesis and Mechanims Like PPY, PANI is prepared by chemical or electrochemical oxidative polymerization of aniline ($C_6H_5NH_2$, MW = 93.13, solubility in water = 36 g/L at 20 °C). Other techniques involve photochemically initiated polymerization [21,22], and enzyme-catalyzed polymerization [23,24]. EP of aniline is performed in an acidic electrolyte (pH 1–3) as short conjugation oligomeric materials with different properties are formed at higher pH [25]. The first anodic oxidation is rate-determining and irreversible to produce cation radicals [26]. The formation of radicals is confirmed by introducing resorcinol, hydroquinone, benzoquinone, etc. to retard and even terminate the reaction [27]. The subsequent steps are highly dependent upon the reaction conditions if two anilinium radicals at the electrode surface form a dimer (Scheme 2). The stable radical cations react to form low molecular weight products that are soluble with aniline, whereas the unstable cation radicals react very fast with anions or solvent molecules to form undesirable side products [28]. The chain propagation step involves the simultaneous oxidation of the dimer and aniline on the anode followed by the coupling of oligomer radical cations with anilinium radical cation. Finally, anion from the dopant adds to the polymer backbone to form the conducting polymer (Scheme 2).

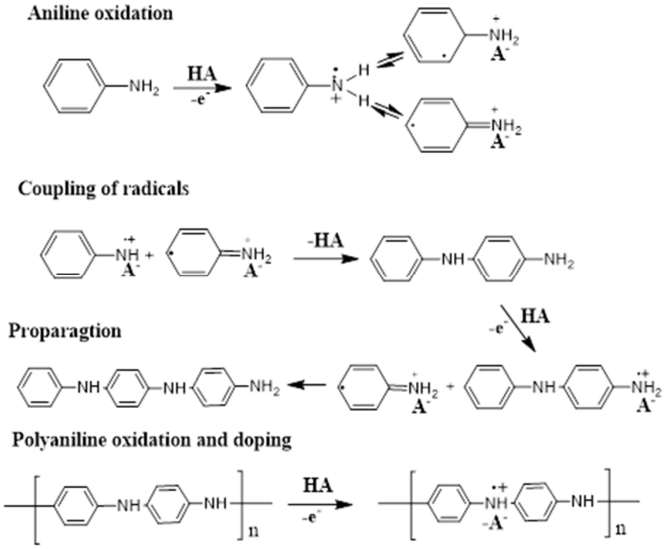
The application of polyaniline in biochemical systems might be limited as this polymer losses its activity at pH above 4 [29,30]. The deposition of aniline (0.1–1.0 M) is commonly performed under sulfuric acid [31] or hydrochloric acid [32]. Aniline can also be electropolymerized in aqueous electrolytes, acetonitrile, and ionic liquids [11,25]. The introduction of "so-called" pH functional groups such as sulfo, carboxyl, or hydroxyl has been attempted [30]. Aniline can be co-polymerized with *o*-aminobenzoic acid, *m*-aminobenzoic acid, or *m*-aminobenzensulfonic acid to gain redox activity at high pH [25]. PANI can also be used with CNTs [33] and other nanomaterials for electrode modification as discussed later. PANI is the first commercial CP, and some sulfonated aniline derivatives are also commercially available. This polymer serves well as a matrix for the immobilization of biomolecules and might function as a mediator for electron transfer in some enzymatic reactions. PANI can be synthesized in the form of varying nanostructures like nanospheres, nanowires, nanorods, and nanotubes. Besides glucose, PANI has been employed in various chemosensing and biosensing platforms, e.g., in the detection of cancer biomarkers, bacteria, infectious diseases, heavy metals, drugs, phenolics, pesticides, sulfites, acrylamides, etc. [34]. Like other conducting and non-conducting polymers, composites of PANI with other CPs, metallic

nanoparticles, graphene, CNTs, etc. have increased surface area, conductivity, and stability. Such desired features play an important role in the development of analytical platforms for a variety of analytes valuable in the food industry, medicine, and environmental monitoring.

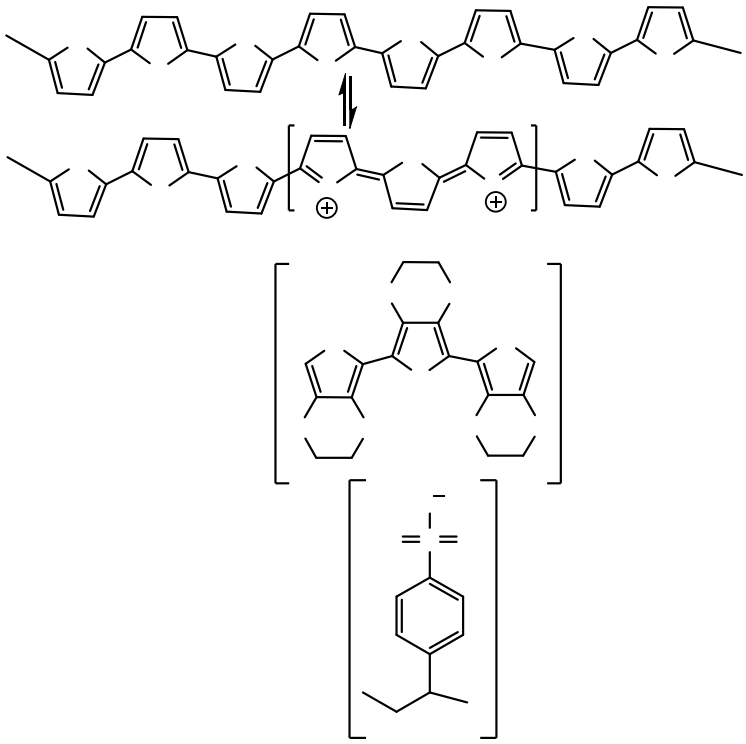
Synthesis of Polythiophene (PTP) and Its Analogs. With an electron-rich aromatic ring, thiophene is oxidized to form well-adhering polymeric films with high conductivity. The film thickness can be controlled by changing the polymerization time, and the structures of PTP and its analogs, as shown in Scheme 3. Of notice for bio-applications is a one-step enzyme entrapment by the EP process based on polythiophenes bearing oligo (ethylene glycol) spacers and porphyrin units [35]. 2,2′-bithiophene with lower oxidation potential, compared to thiophene, can be a good monomer. PTP and substituted PTP are electrochemically deposited as very adhesive thin films or as thick powdery deposits [36]. Poly(3,4-ethylene dioxythiophene)-polystyrene sulfonate (PEDOT: PSS) is a polymer mixture or macromolecular salt. PEDOT is a conjugated polymer with a positive charge, whereas PSS carries a negative charge due to the presence of deprotonated sulfonyl groups. There are four water-soluble polythiophene derivatives, which are synthesized via oxidative copolymerization by $FeCl_3$ [37].



Scheme 1. (**1**) Simple structural representation of polypyrrole (PPY); (**2**–**3**) overall reaction of polymerization and doping, and (**4**) poly(2,5-di-(-2-thienyl)-pyrrole.



Scheme 2. Synthesis steps of polyaniline (PANI) and its resonance structures.



Scheme 3. (**1**) polythiophene (PTP), (**2**) poly(3,4-ethylene dioxythiophene) (PEDOT), and (**3**) polystyrene sulfonate (PSS).

3. Fundamental Aspects of Electrochemistry and Biorecognition Elements

As the oldest analytical technique, electrochemistry offers a wide range of application possibilities by probing the charge transfer between an electrode and a liquid (electrolyte) or solid phase. In electrochemical sensing, an electrode serves as the transducer element to provide the electron for signal acquisition. This approach offers several advantages such as low cost, fast analysis, and ease of miniaturization. Associated with a biorecognition element, electrochemical sensors, known as biosensors, become specific for a target analyte, as exemplified by the selective detection of glucose using glucose oxidase or glucose dehydrogenase [38].

A typical electrochemical cell consists of three-electrodes: detecting (working), reference electrode, and counter (auxiliary). The use of the counter electrode, normally a large platinum wire with a small charge-transfer resistance, overcomes the potential drop due to the current flowing through the electrochemical cell. This configuration is needed for solutions with poor electrical conductivity. The measurement can be reduced to two-electrodes with a potentiostat that connects the working electrode with a non-polarizable reference electrode, e.g., Ag/AgCl (3 M KCl). Electrochemical techniques can be classified into five categories: potentiometry, amperometry, voltammetry, conductometry, and impedance spectroscopy. Potentiometry is based on the measurement of voltage between a detecting electrode and a reference electrode, e.g., pH electrodes and other ion-selective electrodes. Amperometry relies on a redox reaction at a detecting electrode in the presence of a target analyte, which results in the variation of an electrical signal. This method has been used widely in biosensor construction for a broad range of applications. Amperometry encompasses several techniques such as voltamperometry, cyclic voltammetry, differential pulse voltammetry, square wave voltammetry, linear voltammetry, etc.

In brief, cyclic voltammetry is first performed to establish the oxidation/reduction potentials of analytes and assess whether the redox process is reversible, pseudo-reversible, or irreversible [39]. Voltamperometry (current response at a fixed potential) is then conducted to estimate the limit of

detection (LOD), linearity, detection sensitivity, stability of the biorecognition element, and reproducibility for a given biosensor. Differential pulse or square wave voltammetry is useful for analyzing multiple analytes or endogenous reactive species in the same sample [40]. Conductometric relies on the measurement of conductivity or resistivity [41]. A capacitor changes its capacitance when exposed to the desired analyte, which binds to the capacitor's surface. In impedance spectroscopy (EIS) or impedimetry, an alternative voltage (V_{ac}) is applied with frequencies ranging from 10^{-3} to 10^{5} Hz. The resulting current at the same frequency (linear system) is measured. The resulting impedance (voltage/current) is represented on a Nyquist diagram, with the imaginary part as a function of the real part. Impedance spectroscopy has been used to probe cell behavior at an electrode surface in the presence of organic compounds, drugs, nanomaterials, etc., [42,43]. Commercial electrochemical workstations are well equipped with powerful software for conducting electrochemical experiments with different techniques. This subject has been well described in several textbooks [44] and review papers concerning electrochemical sensors/biosensors [45].

Biorecognition elements can be enzymes, DNA, peptides, aptamers, whole cells, etc. and the immobilization of such biomolecules on different types of electrodes has been available elsewhere [46]. Other special cases are the immobilization of glucose oxidase on an organic transistor using a polymer brush [47] and the hybridization of DNA on conducting polymer film-modified diamond [48]. Different biosensing formats have been attempted including the fabrication of an HRP (horseradish peroxidase)-based biosensor by thermal inkjet printing [49],a glucose sensor by a nanocomposite [50], and miniaturization technologies for biosensor design [51]. The potential of continuous measurement of analytes by biosensors in biological media is also reviewed [52].

4. Retention of Biorecognition Molecules by Conducting Polymers-Electropolymerized Polymers/Monomers

A single-step electrochemical immobilization method is used for the production of enzyme entrapped PANI thin films for glucose and urea sensors [53]. Both enzymes and PANI are deposited simultaneously from an aqueous solution of aniline containing optimized amounts of glucose oxidase and urease for the fabrication of two different biosensors. The chronopotentiometry curves for both the biosensors differ mainly in the total time taken for deposition. Other parameters, such as the type and enzyme concentration, have a negligible effect on the curve shape. The one-step immobilization technique has helped to produce functionalized films possessing an optimal working pH close to that of free enzymes to conserve their conformation and microenvironment. PPY, PEDOT: PSS, and polyaminobenzoic acid (PABA) are some common polymers used in biosensor fabrication [54–57]. A biosensor for 17β-estradiol detection is fabricated based on a benzothiadiazole derivative of EDOT [58]. The use of this polymer leads to the formation of a granular defect-free surface structure with a pore diameter of 1 μm. This porous structure allows enzyme molecules to anchor freely in the available space with no covalent bonding. The fabricated surface has a high enzyme loading capacity and easy regeneration capability where the sensor can be regenerated just by washing with deionized water. A new semiconducting polymer poly(9 -nonyl-2,7-di(selenophen-2-yl)-9*H*-carbazole) is used in designing biosensors for norepinephrine determination [59]. Laccase is immobilized on the surface by π-π stacking involving the interaction of π-bonds in the enzyme structure with π-bonds of the aromatic backbone. The interaction stabilized the tertiary protein structure, thus improving its stability and activity. In addition to providing a surface for biomolecule immobilization, the presence of a small amount of selenium in the polymer has a positive effect on laccase catalysis.

EP has been proved as a simple and effective method to deposit conducting polymers on the conductive surface. The property of conducting polymers can be manipulated just by changing the composition of electrolytes used for polymerization. EP is performed to deposit zwitterionic PEDOT conducting films with antifouling properties on a conducting substrate. A mixture of thiophene-based compounds bearing phosphorylcholine (poly(2-methacryloyloxyethyl phosphorylcholine)), carboxybetaine (poly(carboxy betaine methacrylate)), and sulfobetaine (poly(sulfobetaine methacrylate) groups enhances the anti-biofouling properties [60]. The surfaces show resistance to protein and cell adhesion, which would help in improving the shelf life of the biosensors. An

antifouling zwitterionic polypeptide modified with an additional linker of four residues is prepared and used for cancer biomarker BRCA1 (BReast CAncer gene) detection [61]. The composite shows strong resistance to non-specific protein adsorption and enhanced stability. Molecularly imprinted polymers (MIPs) based sensors are low cost and simple to fabricate. A typical MIP would include functional monomers, crosslinking reagents, and template molecules. As the polymer template is removed, a microcavity is formed for specifically binding target molecules with the same shape and structure as a key and lock [62]. The overuse of antibiotics is a major public health concern as they increase the chances of being antibiotic-resistant. It also harms the environment because of the pharmaceutical effluent mixing in water bodies. Amoxicillin (AMO), is one of the most popular types of aminopenicillin. It is detected by a hybrid film of (poly(methacrylamide-vinyltrimethoxysilane (VTMOS)-tetraethoxysilane (TEOS)/AMO) spin-coated on a surface plasmon resonance (SPR) gold sensor. The fabricated sensor detects the AMO down to 73 pM and differentiates among similar molecules, both in the buffer and in water [63]. Artemisinin is also a well-known and widely used drug against malaria, a deadly disease in tropical regions of the world. Ortho-phenylenediamine (*o*-PD) is electropolymerized in the presence of artemisinin on the gold wire. It provides highly stable and effective binding cavities for the target molecule with a detection limit of 0.01 μM [64]. However, it is still unclear how the interaction of the target molecule with functional monomers in the polymer can influence the microcavity formation. Some proteins can easily react with the polymer and change the tertiary and quaternary structure of the template, thus lowering the sensor selectivity.

For controlling the properties of CPs, different strategies have been used, and one of them is chirality. Chiral CPs with shown diverse structures enhance electronic properties and also help in redox reaction and pH switching [65]. They can help to understand bio-related processes due to their remarkable molecular recognition capability and distinctive self-assembling properties. The fabrication of chiral sensors is a promising field of research for electrochemists. A chiral supramolecular polymer consisting of alanine-based naphthalene diimides and Zn ions (AlaNDI-Zn) [66] exhibits multifunctional properties. Its photoconductive and chemo-resistive characteristics can be used in developing various sensors. It detects toxic hydrazine with a detection limit of 3.2 μM.

Hydrogels are polymers with high water retention capacity resembling biological tissues, which show great potential in biomedical applications [67,68]. Various target responsive hydrogels have been developed for molecular diagnostics, and wearable biosensors. A hydrogel-based localized surface plasmon resonance (LSPR) sensor is useful for detecting the biomarker in human tears. Silica gold nanoshells with poly (*N*-isopropyl acrylamide-*co*-methacrylic acid) (PNM) hydrogels can be used to detect lysozyme and lactoferrin. It shows potential as an affordable screening tool for chronic dry eye conditions [69]. A multiplexed biosensor based on PANI hydrogel is fabricated using an inkjet printer. The printing process consists of three steps with a multi-nozzle inkjet printer. The PANI hydrogel is printed, followed by the deposition of platinum nanoparticles, and enzymes. The multiplexed sensor can detect glucose, lactate, and triglycerides simultaneously with high selectivity in real-time [70]. A list of different polymers used for the electrochemical biosensors along with the method of biomolecule immobilization and other sensor properties are given in Table 1.

Table 1. The use of polymers to retain the recognition molecules or modification of the electrode surface.

Polymer (s)	Analyte	Biomolecule Immobilized	Immobilization Method	LOD	LDR	Ref
PANI	Glucose (G) and urea (U)	GOX/URE	Entrapment	-	0.1 mM (G) 10^{-5}–0.1 M (U)	[53]
PEDOT	Lactic acid	LDG	Covalent	-	0.05–1.8 mM	[71]
PBETB	17β-estradiol	HRP	Covalent	10^5 nM	0.1 to 200 mM	[58]
PEDOT/PSS	H_2O_2	HRP	Adsorption	0.4–10 mM	9.4 nM	[72]
PPY-pTS	Hypoxanthine	XOX, URI	Entrapment	5 μM	5 μM–5 mM	[54]
PPY	Interleukin-10	AIA	Covalent	0.347 pg/mL	1–10 ng/L	[55]
PPY/PSS	Glucose	GOX	Entrapment	-	0.1 nM–1 mM	[73]
NDSC	Norepinephrine	Laccase	Adsorption	240 nM	0.1–200 μM	[59]
DMAPM, VPBA, EGDM, DA	Dopamine	Imprint	Entrapment	96 nM	40 nM–20 μM	[74]
MAA, EDGM, fluorescein-o-acrylate	Doxycycline	Imprint	Entrapment	117 nM	0.2 to 6 μM	[75]
Toluidine blue	CA 15–3	Imprint	Entrapment	0.10 U/ mL	0.10 U/mL to 100 U/mL	[76]
PDA	TNT	Imprint	Entrapment	15.4 nM, 14.8 nM	0–289 nM	[77]
o-DB, o-PD	Chlorpyrifos	Aptamer imprint	Entrapment	0.35 fM	0.001 to 0.4 pM	[78]
Macroporous PANI, PSS	AFP	AFP Ab	Covalent	3.7 fg/mL	0.01 to 1000 pg/mL	[79]
CuPT-PPy/NIPAAm-AM	Melanoma	CRP	Covalent	7.85×10^{-19} M	10^{-20} to 10^{-6} M	[80]
NiPAAm, AAm, MBAAm	BSA	Imprint	Entrapment	0.012 μM	0.02 μM to 10 μM	[81]
FcMMA, 4-VP, and EDMA	Bisphenol A	Imprint	Entrapment	13 ng/L	0–410 ng/L	[82]
PEDOT	NADH		Physical adsorption	5.3 μM	20–240 μM	[83]
PLL	DNA hybridization	DNA	Covalent	2.25×10^{-14} M	1.0×10^{-13} to 1.0×10^{-6} M	[84]
PPY	CEA	Imprint	Entrapment	0.13 pg/mL	0.125–12.25 pg/mL	[85]
PCPy	Glucose	GOX	Entrapment	0.039 mM	0.1–15.0 mM	[86]
PSBEDOT	Glucose	GOX	Entrapment	-	1 mM to 20 mM	[87]
BTP	Glucose	GOX	Entrapment	0.034 mM	0.034–1.0 mM	[88]
PPY	Glucose	GOX of *A. niger*	Coating the cell	0.005 mM	0.01–0.05 mM	[89]
PPY	Catechol	POX (*Trametes pubescens*)	Encapsulation	1.8 μM	1–60/70 μM	[90]
p-toluene-sulfonic acid doped PPY	Glucose	GOX	Physical adsorption	0.25 mM is lowest tested	0.25–1.5 mM	[91]
BOTT	Catechol	Laccase	Entrapment	0.38 μM	0.5–25 μM	[92]

PEDOT(BSA)	Glucose	GOX	Physical adsorption	0.416 mM is tested	0.416 to 50 mM(LSV)	[93]
DPDAA	Glucose	*G. oxydans* and GOX	Glutaraldehyde cross-linking Cells:	0.022 mM (GOX) 0.081 mM (whole cell)	0.045–50.0 mM (GOX) and 0.19–50.0 mM (whole cell)	[94]
PANI	*E. coli*	Anti-*E. coli* antibody	Covalent	10 CFU/mL	10^1–10^6 CFU/mL	[95]
Macroporous PPY-pyrrole	Urea	Urease	Entrapment	2.57 mM	1.67–8.32 mM	[96]
PEDOT	BRCA1	DNA	Covalent	0.0034 pM	0.01 pM to 1 nM	[61]

Polymers: AAm = acrylamide, BOTT = thienothiophene-benzoxadiazole-alt-benzodithiophene, BTP = 5-amino-N1,N3- bis (2,5-di(thiophen-2-yl)-1*H*-pyrrol-1-yl) isophthalamide, DA = 3-hydroxytyramine hydrochloride, DMAPM = *N*-[3-(dimethylamino)propyl]methacrylamide, DPDAA = 10-(4*H*-dithiyeno [3,2-b:2′,3′-d]pyroll-4-il)decan-1-amine, EGDM = ethylene glycol dimethacrylate, FcMMA= Ferrocenyl methylmethacrylate, MAA = methacrylic acid, MBAAm = methylene-*N*,*N*-bis(acrylamide), NDSC = 9-nonyl-2,7-di(selenophen-2-yl)-9*H*-carbazole, NiPAAm = N-isopropylacrylamide, o-DB = o-Dihydroxybenzene, PANI = polyaniline, PBETB: poly(4,7-bis(5-(3,4-ethylenedioxythiophene)thiophen-2-yl)benzothiadiazole), PEDOT = poly(3,4-ethylenedioxythio phene), pTS = para toluene sulphonate, PLL = Poly-L-lysine, PPy = poly pyrrole, PSBEDOT = poly(sulfobetaine-3,4-ethylenedioxythiophene), PSS = poly styrene sulphonate, VPBA= 4-vinylphenylboronic acid, 4-VP= 4-vinylpyridine. *Biorecognition biomolecules*: AFP = alpha-fetoprotein, AIA = Anti-Interleukein antibody, BSA= bovine serum albumin, CA 15-3 = Carbohydrate Antigen 15-3, CEA = Carcinogenic embryonic antigen, CRP = C-reactive protein, GOX= glucose oxidase, HRP = horseradish dehydrogenase, LDH = lactate dehydrogenase, POX = polyphenol oxidase, TnT = Cardiac Troponin T, URA = urease, URI = uricase, XOX = xanthine oxidase.

5. Conducting Polymer Composites

The properties of polymers can be improved by the addition of fillers and the formation of composites. The most used materials for polymer composite formation are CNTs, carbon nanofibers, graphene and its derivatives, other carbon materials, metal nanoparticles, and metal oxide nanoparticles. CNTs improve the electrical and thermo-mechanical properties of CPs thin films. In general, composites of polymers with both multiwalled CNTs (MWCNTs) and single-walled CNTs (SWCNTs) have been used for several analytes with enhanced sensitivity and stability due to increased and faster biomolecule to electrode charge transfer as well as improved biomolecule activity. Graphene is another carbon-based nanomaterial that has been used widely in electronic devices. A field effect transistor (FET) based on a composite of graphene and poly(3-aminobenzylamine-*co*-aniline) exhibits extremely sensitive detection of acetylcholine. The amine-functionalized polymer provides a suitable environment for enzyme immobilization and improves the pH sensitivity of the graphene-based FET owing to a wide pK_a distribution of polymer components [97]. An electrochemical comparison between GO/PEDOT and rGO/PEDOT composite reveals that GO enhances the capacitive behavior of the polymer film more than rGO [98]. Currently, porous carbon with a high conductivity and surface area has emerged as a new cost-effective material for screen printed electrodes used in biosensing devices. Its electrocatalytic properties can be further enhanced by doping carbon with heteroatoms like B (boron), N (nitrogen) or S (sulfur). The detection of neurotransmitters and lactic acid is feasible using the composites of N and S doped carbon with polymers like poly 3-((2,2′:5′,2′′-terthiophen)-3′-yl)-5-aminobenzoic acid (pTTABA)) and 2, 2′:5′, 5″-terthiophene-3′-p-benzoic acid (TBA) [99,100]. The core-shell nanocomposite of Pt-Pd with N-doped carbon shell acetylcholine biosensor is fabricated for organophosphate pesticide detection in "real-world" samples. The composites of porous carbon with polymers show better electrocatalytic properties and conductivity. Its detection limit for malathion, chlorpyrifos, and parathion is 7.9×10^{-15} M, 7.1×10^{-14} M, and 8.6×10^{-15} M, respectively [101]. Carbon dots (CDs) are the fluorescent nanosized carbon particles which have attracted much attention since their discovery. In comparison to other carbon materials, CDs are easy to synthesize, small, hydrophilic, and can be modified with various functional groups. CDs with fluorescent and photophysical properties are often mixed with polymers to enhance the electron transport and stimulate fast redox reactions. CDs are mixed with chitosan for the determination of glucose as the composite enhances the electron transfer and provides a biocompatible sensor surface. The CDs increased the conductivity and 3-aminobenzeneboronic acid as the monomer for the fabrication of MIP. The fabricated sensor exhibits a limit of detection (LOD) of 0.09 μM with excellent stability, and reproducibility [102]. Caffeic acid (3,4-dihydroxycinnamic) is a widely used drug for anti-inflammatory, leucopenia, and thrombocytopenia, and hemostasis. A CDs@MIPs based sensor detects this acid in human plasma with an LOD of 0.11 μM and a linear range of 0.5–200 μM [102].

Graphitic carbon nitride (gCN) is a synthetic polymer with an increase in interest in applications ranging from photoelectrochemistry, (photo)catalysis, and, biosensors [103]. Its reactive interface with an aromatic π-conjugated framework provides opportunities to control its structure and properties for a specific application. The gCN is incorporated with MIP to fabricate a self-powered biosensor based on photo-fuel cell, resulting in good detection selectivity for bisphenol A. The sensor is powered through the catalytic activity of $Ni(OH)_2$ toward the oxidation of glucose [104]. An electroactive nanocomposite of Ru@Au core-shell nanoparticles, gCN nanotubes, and functionalized graphene quantum dots is used to detect phenylethanolamine A [103]. The fabricated sensor with an LOD of 2.0×10^{-13} M is tested on the urine sample with good selectivity for phenylethanolamine A. Gold nanoparticles (AuNPs) are the most used metal nanoparticles in biosensors. An interesting antifouling biosensing platform is constructed using AuNPs and PEDOT, where AuNPs acted as signal enhancers, whereas the polymer assisted in the antifouling behavior of the sensor surface [105]. The device is tested in biological media for CA15-3 detection, and it shows an elongated antifouling performance with only a negligible activity change during 30 days. Natural clay has been under research as a transducer surface modification agent due to its stability, porosity, high cation exchange

capacity, low price, and widespread availability [106,107]. A glucose biosensor based on a platinum disk modified with PANI, chitosan, montmorillonite clay particles, and platinum nanoparticles shows extended stability up to two months with excellent stability and selectivity [108]. Thus, clay particles can be used to increase the stability of the polymer surface-modified electrode.

Table 2. Polymer nanocomposites containing carbon nanotubes (CNTs), graphene, and nanoparticles in biosensors.

Nanocomposite	Target Analyte	LOD	LDR	Ref.
PEDOT and iridium (IV) oxide	Catechol (C) and azinphos methyl (AM)	0.017 μM (C) 2.964 μM AM	0.05–10.65 μM C 10–200 μM AM	[109]
PDTP and MWCNTs	Glucose	2.2–0.43 μM	0.43 μM	[110]
PANI, Pt, MMT, and CS	Glucose	10 μM–1.94 mM	1.40 μM	[96]
PABA and Graphene	Acetylcholine	5 to 1000 μM	2.3 μM	[97]
PEDOT and AuNPs	CA15-3	0.001–1000 U mL-1	0.32 mU/ mL	[105]
PTAB and N, S doped porous carbon	Lactic acid	0.5 μM and 4.0 mM	112 nM	[99]
MnO_2/PAA-GG	H_2O_2	10 μM	0.05 mM–50 mM	[111]
PEG/graphene/aptamer	IgG	47 pM	50 pM–250 nM	[112]
PQQ-GDH/PTh/MWCNT	Glucose	1 μM	1 μM to 2 mM	[113]
PANI/MWCNTs/Starch	Cholesterol	0.01 mM	0.032 to 5 mM	[114]
GOX-PtNPs-PEDOT-MS	Glucose	1.55 μM	0.1–10 mM	[115]
CeO_2-NRs/Ppy-NPs ssDNA probe (covalent)	DNA from *Salmonella*	0.29 μM	1.0 nM–1 μM	[116]
PIn-5-COOH/MWCNTs-COOH	α-fetoprotein	0.33 pg/mL	0.001–100 ng/mL	[117]
poly(pyrrole-3-carboxylic acid)/GO with GOx	Glucose (ITO coated glass)	0.05 mM	1–20 mM	[118]
PAnNFs-P(An-co-PoPD)]	Glucose (GOX)	Minimal glucose conc. tested is 5 μM	Saturation at 0.45–0.6 mM	[119]
PDA-MWCNTs -PB	Cholesterol (Ch) (Ch oxidase/ SPE)	1.5 μM	Linear up to 0.4 mM	[120]
PEG-PPY nanowires	DNA probe for miRNA	0.033 pM	0.10 pM–1.0 nM	[121]
rGO-PEDOT:PSS	GOX fpr glucose	86.8 μM	Detection range: 100 μM to 600 μM	[122]
PPY-Ag/ZnO on pencil graphite electrode	Xanthine (X) (X oxidase)	0.07 μM	0.06–0.6 μM	[123]
P(SNS-An-co-EDOT)	Glucose (GOX)	1.9 μM	0.01–5.0 mM	[124]
PPY-CNTs/(SrCuO2)	2,4-dichlorophenol (laccase-graphite electrode)	0.18	1–50 μM	[125]
PPY bearing the redox dendrimer PAMAM G2	autocrine motility factor	43 fM	1 pM–1 μM	[126]

PANI/PPY/MWCNTs-COOH	(cytokine)-cancer biomarker H_2O_2 (Cytochrome C)	0.1 μM	1–370 μM	[127]
ePDA	AA (Laccase)	The lowest AA tested is 7 μM (CV mode)	Sensitivity = 342 mA/Mcm^2 with a linearity of 1–25 μM	[128]
Pt/rGO/P3ABA	-Glucose (GOX) -Cholesterol (ChOx)	44.3 μM 40.5 μM	0.25–6.00 mM 0.25–4.00 mM,	[129]
PtNPs, MWCNTs-COOH – PEDOT:PSS	Glutamate Glutamate Oxidase	0.5 μM at 0.5 V 0.5 μM at −0.2 V	1 μM and 800 μM 10 μM and 600 μM	[130]
Cu-BTC- PANI on ITO	*E. coli* (anti-*E.coli* antibody)	*E. coli* (2 cfu/mL)	Linearity: 2–2 × 10^8 cfu/mL	[131]
PBA-pTBA	Glucose (GOX)	0.29 μM	1 μM–30 mM	[132]
PDA	ABTS (Laccase)	0.29 μM	1–150 μM	[133]
PANI/rGO/Au NPs	Glucose GOX	64 μM	1–10 μM	[134]
PAB/ DNA-functionalized CNT/Nitrogen doped graphene	Dopamine	14 nM	0.02–1 μM	[135]
PANI/ Au NPs	Prostate-specific antigen Peptides	0.085 pg/ mL	0.1 pg/mL to 100 ng/mL	[136]
PAA/VS-PANI/GPL-FePc-conducting hydrogel	Glucose GOX	6.4 μM	1–20 mM	[137]
MPNFs of SnO_2/PANI	Glucose GOx/HRP	1.8 μM	5–100 μM	[138]
PANI/Polyacrylamide	H_2O_2	2.9 nM	0.01–50 μM	[139]
PPY	*M. tuberculosis*	0.36 aM	0.1 aM to 100 fM	[140]
Carbon black/PVDF/SPGMA	IL-8	3.3 fg/mL	0.01–3 pg/mL	[141]
PEDOT: PSS/Fe_2O_3	Carcinoembryonic antigen	-	4–25 ng/mL	[56]
AChE/HCS@PANI	malathion	1.0 ng/mL to 10 μg/mL	0.16 ng/mL	[142]
GO/AAM/ MAA/MMA/ NVP	Zika virus	10^{-3} to 10^2 PFU/mL	2 × 10^{-2} PFU/mL	[143]

AA = 2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid, ABTS = 2,2′-azino-bis-(3-ethylbenzothiazoline-6-sulphonic acid), AChE = acetylcholine esteras, BTC = 1,3,5-benzenetricarboxylic acid, CMA = 4-aminophenylacetic acid, CS = chitosan, PABA = poly(3-aminobenzylamine-co-aniline), ePDA = a co-polymer enriched by chains of dopamine in the open form in detriment of indole moieties, GPL-FePc = iron phthalocyanine functionalised graphene nanoplatelets, HCS = hollow carbon spheres, HRP = horse radish peroxidase, MMT = montmorillonite, MPNFs = Polymerized multiporous nanofiber, P(An-co-PoPD)] = poly(aniline-co-N-phenyl-o-phenylenediamine) nanoflakes, P(SNS-An-co-EDOT) = poly(benzenamine-2,5-di(thienyl)pyrrole-co-3,4-ethylenedioxythiophene, P3ABA = poly(3-aminobenzoic acid, PAA = polyacrylic acid, PAB = poly anilineboronic acid, PAMAM G2 = N-(5-phosphate-D-arabinoyl)-2-aminoethanamine, PAnNFs=

Polyaniline nanofibers, PB = Prussian Blue, PBA = poly-3′(benzoic acid),PDA = polydopamine, PDTP = poly 2,5-di(thienyl)pyrroles, PEDOT = poly(3,4-ethylenedioxythiophene), PEG = polyethylene glycol, PIn-5-COOH = poly(indole-5-carboxylic acid), PPY = poly pyrrole, PQQ-GDH = pyrroloquinoline quinone dependent glucose dehydrogenase, PTAB = polymer (poly 3-(((2,2′:5′,2′′-terthiophen)-3′-yl)-5-aminobenzoic acid, pTBA = poly(2,2′:5′,2′- terthiophene), PTh = polythiophene, PVDF= polyvinylidene fluoride, SPGMA = star poly(glycidyl methacrylate), TBA = 5″-terthiophene-3′-p-benzoic acid, VS-PANI = inyl substituted polyaniline.

As illustrated in Table 2, conducting polymer-based composites are formed by blending or mixing CPs with other materials to achieve multifunctionality [144]. Resulting composites benefit from the synergistic effect of different components to exhibit multi-functionalized properties with enhanced mechanical performance and processability. The interfacial adhesion between the conducting polymer and other components affects the morphologies and electrochemical properties and morphologies of composite films, which can be attributed to novel composite properties. The presence of the amino group of aniline is needed for bioconjugation with the amino group or the carboxyl group of the biorecognition element by glutaraldehyde crosslinking or carbodiimide activation, respectively. Noble metal nanoclusters, e.g., platinum nanoclusters, can be integrated during oxidative polymerization of pyrrole to form a composite with enhanced electrochemical properties, mechanical performance, and processability [145]. The conducting polymer-based composite exhibits electrocatalysis for glucose oxidation with a linear range of 1–30 mM. However, the signal response is not provoked by common electroactive interferents such as ascorbic acid, uric acid, and 4-acetamidophenol in blood glucose analysis.

6. Polymers Conjugated Mediators for Direct Electron Transfer (Det)

As third generation biosensors are mainly based on DET from the enzyme–electrode surface, the polymer matrix can serve as an efficient surface for mediator attachment, thus preventing the leaching of mediator molecules and improving the electron transfer as well. Ferrocene and its derivatives are very popular redox mediators used in biosensor construction. They are very stable in both oxidized and reduced form, have excellent redox reversibility, and their redox capacity is not affected by the pH of the solution [110,140]. A highly sensitive reagentless glucose detection electrode is designed with copolymerization of ferrocene and functionalized PPY. The sensor has a detection limit of 0.43 μM with a sensitivity of 23.12 μA mM^{-1} cm^{-2} and can detect trace analytes [110]. The functionalization with the mediator shifts the working potential from 0.6 V to 0.4 V and prevents mediator leaching, thus increasing stability and limiting the interference in real samples. A biosensor surface is modified with ferrocene and PPY for bacterial detection with a LOD of 100 fM [140]. A sensitive microfluidic set-up coupled with a modified SPE is designed for neurotransmitter detection and able to detect the target analyte as low as 34 pM [100]. Titanium yellow, an aromatic anionic compound, can be used to enhance PPY electroactive properties. The polymeric film with sulfonate functional groups provides enhanced conductivity and stability compared to chlorine [146].

Evans blue acting as a mediator is attached to the polymer surface by covalent linkage. It acts as an electron mediator and aids in selectivity enhancement by attracting cationic species and repelling anionic counterparts due to the presence of negatively charged sulfonate groups. The device provides a rapid, low cost and effective method for neurotransmitter detection and avoids the use of specific labels. Viologen is a well-known redox molecule containing a 4,4′-dipyridinium group. It undergoes two-electron transfer processes and exhibits three redox states, namely V^{2+}, V^{+}, and V^{0}, making it suitable for electrochemical applications. This property of viologen is used to modify carbazole monomer by electropolymerization, which exhibits enhanced thermal stability and a large electrochemical window [131,147]. The polymer-based electrode enhances the oxidative current in the detection of uric acid and ascorbic acid. However, the mediator modification does not always increase the sensor performance. Nickel hexacyanoferrate (NiHCF) is used as a mediator for layer by layer EP of PPY-PSSA and PPY-GOX or lactate oxidase. The sensor without a mediator has higher

sensitivity of 0.94 nA/mM for glucose than the mediator modified sensor (0.32 nA/mM) [148]. A list of various mediator conjugated polymer platforms used in 2020 for biosensing is given in Table 3.

Table 3. Biosensors consisting of mediator conjugated polymers for direct electron transfer.

Mediator-Polymer	Target Analyte/ Sensor Platform	LOD	LDR	Ref.
Evans blue-poly-TBA/S,N doped carbon	Neurotransmitters (Amp)	0.034 (±0.005) nM 0.044 (±0.004) nM	0.05−130 nM	[100]
Ferrocene-PPY/MWCNT	Glucose (Amp)	2.2–0.43 μM	0.43 μM	[110]
Ferrocene-PPY	*M. tuberculosis* (SWV)	1 aM to 100 fM	0.36 aM	[140]
Polymethylene blue-PANI/MWCNT	Cardiac troponin T (DPV)	0.10–8.0 pg/mL	0.040 pg/mL	[149]
Ferrocene-PAA hydrogel	miR-21 (DPV)	10 nM to 50 μM	5 nM	[150]
BG-PAA	Glucose (Amp)	0.25–5.0 mM 0.1–1.75 mM	0.055 mM	[151]
FcAPS/GCE	H_2O_2 (Amp/Fluors)	10 μM to 10 mM	2.07 μM	[152]
FPS/MWCNT- PDDA	Acetaminophen (DPV)	3 to 1100 μM	0.6 μM	[153]
PEDOT/AuNPs- MWCNT-COOH	Catechol Laccase (DPV)	0.11 and 12.26 μM	Two linear ranges: 0.1–0.5 and 11.99–94.11 M	[154]

Amp = Amperometry, BG = Brilliant Green, FcAPS = ferrocene carboxylic acid (FcA) cross-linked amino-polyethersulfone film, Fluors= Fluorescence, FPS = ferrocene polymer nanospheres, PAA = polyacrylamide, PANI = polyaniline, PDDA = Poly diallyldimethylammonium chloride, DPV = differential pulse voltammetry, PPY = poly pyrrole, TBA = 5″-terthiophene-3′-p-benzoic acid.

7. Non-Enzymatic Detection of Glucose

Table 4. Non-enzymatic glucose detection with conducting polymers (CPs) matrices.

CPs	Metal Particles	Analytical Performance	Ref.
Polybithiophene	Pd microparticles	Linearity: 0.04 to 0.4 mM LOD: 7 μM (under alkaline conditions)	[155]
Keggin Polyoxometalate/MWCNTs	Mixed-valence Co(III) and Co(II) structures	Linearity: 0.1 mM to 10.0 mM LOD: 1.21 μM (in 0.1 M NaOH)	[156]
PPY nanowire	A nickel-cobalt-sulfide nanosheet	Two linear ranges: 2 μM to 140 μM 0.14 mM to 2 mM LOD: 0.82 μM (in 0.1 M NaOH)	[157]

Benzoic acid-functionalized poly(terthiophene)	AuNPs	Potentiometric response (in saliva): 0.32 μM - 1 mM, LOD = 0.19 μM.	[158]
Poly (*o*-phenylenediamide)	Ag-NPs	Linearity: 0.15 to 13 mM LOD: 12 μM Detecting blood glucose, PBS pH 7	[159]
Poly(thieno[3–4-b]-1,4-dioxin-2-methanol)	Not required	Linearity: 1–9 mM In 0.1 M PBS	[160]
PANI nanofiber	NiO/CuO NPs	Linearity: 20 to 2500 μM LOD: 2 μM (in 0.1 M NaOH)	[161]
PPY	$ZnFe_2O_4$ magnetic nanoparticles	Linear up to 8 mM LOD: 0.1 mM, in).1 M NaOH	[162]
PPY nanowires	CuO-Cu_2O NPs	Linear up to 8 mM LOD: 6.2 μM (in 50 mM NaOH)	[163]
Poly(3-octylthiophene)	Au NPs	Potentiometric, Linear: 5–30 mM, LOD: 0.2 mM, in boric acid-borate buffer, pH 9 plus 100 mM NaCl	[164]
PPY nanowire/nickel foam substrate	$NiCo_2O_4$	LOD = 0. 22 μM, linearity: 0.001–20 mM (in 0.1 M NaOH)	[165]
Conductive Polyaniline Nanosheets (CPANINS)		LOD = 0.043 μM, Linearity: 1–1000 μM (in 0.1 M H_2SO_4)	[166]
PANI/MWCNTs	ZnO	LOD: 0.1 mM, two linear regions: 0.1 to 1 mM and 1 to 6 mM (10 mM NaOH)	[167]
PANI	NiO	LOD: 0.06 μM, linearity: 1–3000 μM	[168]

AuNPs = gold nanoparticles, AgNPs = silver nanoparticles, PANI = polyaniline, PPY = polypyrrole, MWCNT = multiwalled carbon nanotubes.

Enzymeless electrochemical determination of glucose has been a dream of many researchers to overcome the drawbacks of poor stability and the high cost of GOX. In this context, considerable attempts for non-enzymatic detection of glucose using CPs, particularly, PPY and PANI, have been noticed and, in most cases, metal or metal oxides are required to effectuate in electrocatalysis (Table 4). The CPs aid in increasing the stability and enhancing the electrochemical capacity of the electrodes. In contrast, biosensors are highly specific and less vulnerable to interference, but the working conditions and stability are still major drawbacks [132]. The measurements are performed in alkaline solution, which is not very attractive to patients with diabetes mellitus, and the linear range is very narrow, falls short of the required limit of 25–30 mM. Non-specific protein binding on metal particles and conducting polymers are expected to foul the detecting surface and affect the signal response. Attempts have been made to improve the conductivity and hence sensitivity using a combination of CPs, metal, or metal oxides with carbon nanotubes or graphene. There are still formidable challenges concerning the applicability of enzymeless detection of glucose using blood [169]. Novel materials offering better electrodes and protective films with antifouling properties

should be investigated to yield mature technologies applicable to the commercialization of non-enzymatic sensors. Non-specific adsorption of proteins and biomolecules is always problematic when more hydrophobic materials are present on the electrode.

8. Trends and Future Possibilities

It has been a long journey since the discovery of CPs with a Nobel Prize awarded to three people: H. Shirakawa, A. G. MacDiarmid, and A. J. Heeger in 2000 [170]. Considerable attempts have been focused on the use of CPs as electrode materials in batteries [171]. There are still several cost and technical issues for the use of CPs in this domain, compared to optoelectronic devices [172]. Electrochemistry has advanced significantly from the pioneering concept and historical use of the mercury electrode by Heyrovsky [173]. A fundamental framework of direct electron transfer (DET) was elucidated and solidified by Marcus, another Nobel Prize winner in Chemistry [174] A search from Web of Science unravels 2500 publications related to the use of CPs for biosensing, starting with the use of two classical CPs, PANI and PPY. Other polymers have emerged and the PEDOT: PSS pair has also received attention, whereas several other CPs have not been attempted [172]. More publications related to CPs have also been supported by nanoparticles, carbon nanotubes, graphene, etc. to form polymer composites. Pertaining to the synthesis of CPs, the polymerization of carbon dots is worth noting [175]. Conducting polymers decorated with nanoparticles such as CuO, zinc doped CuO particles exhibit antimicrobial activities [176,177]. This feature might open the possibility for the development of implanted biosensors or enzyme less detection of glues based on electrocatalysis of such metal oxide nanoparticles.

In principle, biosensors are versatile tools for monitoring of a variety of analytes with significant importance in healthcare, agriculture, environment, food, and biosecurity. The confinement of a recognition molecule on highly conductive electrodes is necessary for the development of sensitive and selective biosensing. Besides glucose monitoring or pregnancy test, the commercialization of other types of biosensors is tepid, including the one for cholesterol. Even glucose biosensors encounter tough competition with non-invasive glucose monitoring systems, equipped with smartphones for data processing, storage, and transmission. Non-specific protein adsorption is always problematic and challenging, which might impede the performance of conducting polymer biosensors. Most of such biosensors have been evaluated under controlled laboratory conditions, and clean samples are spiked with various known interfering molecules. Clinical samples might contain several unknown plausible interferents, which invoke a deficiency in both specificity and selectivity of target molecules. For medical applications, the practice is merely based on single-use, so the issue of simple adsorption on the electrode or covalent coupling of the sensing biomolecule to a conducting polymer becomes very trivial. For environmental monitoring, the degradation of CPs over time is likely unavoidable. Thus, intensive research is still needed on the stability improvement of the biosensors.

The role of existing CPs and the search for new ones need to be addressed from both commercial and academic viewpoints. Pyrrole and its derivatives possess all desirable features for the fabrication of biosensors and immunosensors. In brief, they are stable under ambient conditions and polymerized under natural and aqueous media for the enzyme entrapment process in one single step. Extensive research has been conducted with PPY for the development of chemical sensors, biosensors, DNA-based sensors, and immunosensors.

A similar situation is expected for PANI with an amino group for conjugation of recognition biomolecules as this polymer has been investigated rather extensively in biosensors construction. PANI and its derivative are well-known for their semi-flexibility and high conductivity to serve as immobilization matrices for biomolecules without compromising their biological activities. It can still serve as an immobilization matrix for bioconjugation; however, it must compete with other polymers and biopolymers, including Nafion, chitosan nanocrystal, starch, nylon, cellulose nanocrystal, etc. For long-term applications, pure CPs have several shortcomings such as decreased sensitivity, reduced selectivity, surface poisoning, and other agent intrusions on their pore structures. The issue of polymer aging should be investigated carefully as it might affect the polymer

matrix disorder and permselectivity, resulting in biosensors with poor analytical performances. To a certain extent, CPs can be formulated with graphene, carbon nanotubes, and metallic nanoparticles. Polymer composites could play an important role in the development of wearable and flexible high-performance sensors/biosensors.

The demand for glucose monitoring will inevitably increase significantly as 450 million cases of diabetes could reach 700 million cases in 2045 [178]. The future trend of glucose detection will focus on the different bands of glucose in the electromagnetic spectrum with the aid of advanced mathematical and statistical algorithms to improve the accuracy of glucose estimation [179]. Doubtlessly, the appearance of non-invasive glucose is emerging, and it is just a matter of time before non-invasive monitoring of glucose with acceptable accuracy and specificity is achieved. It is also questionable if continuous monitoring of glucose is needed except for critical diabetes type 1. A comprehensive review of the performance of non-invasive, minimally invasive biosensing for glucose is available elsewhere. Accuracy, selectivity, and precision are always plausible caveats that need to be carefully investigated and solidified, a backdrop for future research endeavors. In emergency departments, clinical diagnosis of glucose, lactate, cholesterol, urea, creatine, and creatinine is performed. Thus, the trend is moving toward the development of integrated biosensors for multi-analytes detection. In this context, the stability of such integrated biosensors is more problematic since this requirement is dependent from one enzyme to another one, which creates significant obstacles to the practical applications. For successful commercialization, the enzymes must be stable for at least 6 months with unique detection selectivity. To date, glucose oxidase is the only enzyme that fulfills these tough requirements, and it is not surprising to see more publications with glucose detection as over 100,000 papers have been published.

The approach of using enzymeless sensing is questionable for any compound due to its lack of selectivity for the analysis of "real world" samples. One example is the detection of nitrite in food, an approved food preservative. Nitrite is a precursor of nitroamines, a suspect cancer agent, so its level is regulated at safety levels [180]. Among various methods, nitrite can be oxidized to nitrate by a gold electrode modified by AuNPs and *p*-aminothiophenol [181]. However, ascorbic acid, an endogenous electroactive species, is also detected to cause significant interference and must be removed by ascorbic acid oxidase if the measurement is based on simple amperometry. Differential pulse voltammetry and square wave voltammetry are more applicable in this situation, provided the potential difference between a target compound and interfering species is sufficient.

The detection of ethanol as an indicator of "driving under influence" is a common practice in many countries. Alcohol oxidase is available, and there have been several attempts with this enzyme for the detection of ethanol in liquid. However, vapor ethanol is well served by different breath analyzers, and this field has been very competitive and crowded with four different approaches: color change due to a chemical reaction, infrared spectroscopy, fuel cell technology, and combined infrared and fuel cell. There has been no significant progress in the respective biomolecule immobilization. Simple adsorption of biomolecules on insoluble supports is attributed to nonspecific hydrophobic interaction and hydrogen bonding. Entrapping biomolecules by polymers, e.g., Nafion is very effective to confine proteins on the electrode. However, bioconjugation is more commonly used to form stable covalent binding between biorecognition molecules with the support for repeated analysis. This technology is maturing, and the most used techniques are still glutaraldehyde and carbodiimide activation.

Based on the solid-phase synthesis, peptide chains with high specificity and versatility can be synthesized and custom-made to meet specific requirements. Compared to other biological materials, peptides are more stable and active for prolonged periods, compared with enzymes and antibodies. Peptides can also be designed to have specific three-dimensional structures like proteins to improve biorecognition capabilities like antibodies and enzymes. The future development of peptide-based biosensors is expected to grow significantly. Considerable demands are also expected for the detection of food and waterborne parasites. *Cryptosporidium*, also known as Crypto (diarrhea linked to water) and *Giardia* (diarrheal illness is known as giardiasis caused by a parasite) are two important targets. Besides human infection, many species of *Cryptosporidium* infect animals, i.e., the

application of biosensors in veterinary medicine could be enormous if it is cost and performance effective, compared to immunoassays. Of importance, is the development of advancing peptide-based biorecognition elements for biosensors using in-silico evolution [182].

Recent progress in the synthesis of porous carbon materials deserves a brief discussion here because the use of these materials in biosensing is very limited. Porous carbon materials have three types of pore sizes: microporous (below 2 nm), mesoporous (>2 nm and <50 nm), and macroporous (above 50 nm) and can be synthesized by several different methods [183]. Porous carbon (PC) has several distinct features such as high surface areas (over 2000 $m^2 g^{-1}$), tunable surface chemistry, and short diffusion pathway for rapid mass transfer. Albeit PC has been used extensively in electrocatalysis, energy storage, and capacitors, its application in biosensors is very limited [184,185] PC can simply form a nanocomposite with Prussian Blue (PB) and serve as an efficient immobilization matrix for GOX toward the development of a biosensor for glucose detection [186]. As electrodes for biosensing, uniform porous carbon materials with highly graphitic structures are needed and the synthesis of such materials has been very challenging. Macroporous graphite felt and porous graphitized carbon monolith are two materials that could be used to fabricate biosensors [187] or high-performance microbial cell anodes [188]. The use of CPs with porous carbon remains to be explored, a backdrop for future research endeavors.

9. Conclusions

Both PANI, PPY, two classical conducting polymers, and their derivatives have been exploited extensively towards the development of different biosensing platforms. Conducting polymer-based composites are formed by blending or mixing CPs with other materials, including metal oxide nanoparticles to achieve multifunctionality. Resulting composites benefit from the synergistic effect of different components to exhibit multi-functionalized properties with enhanced mechanical performance and processability. The interfacial adhesion between the conducting polymer and other components is attributed to composites properties, structures, and morphologies. Noble metal nanoclusters, e.g., platinum nanoclusters, can be integrated during oxidative polymerization of different monomers to form a composite with enhanced electrochemical properties, mechanical performance, and processability. Detection of immunoglobulins such as IgM, IgG, and IgA in response to infectious diseases, will receive considerable attention. Synthesized peptides and recombinant antigens and proteins of viruses and pathogens can be synthesized and used as biorecognition elements. CPs and their composites are anticipated to serve as excellent immobilization matrices for such biomolecules. Emerging porous carbon and carbon felt have potential applications in both chemosensing and biosensing platforms due to their high surface areas and ease of synthesis.

Author Contributions: All four authors wrote the different sections of the paper. J.H.T.L. edited and mainly revised the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: B.D. Malhotra is thankful to the Science and Engineering Research Board (Government of India) for the award of a Distinguished Fellowship (SB/DF/011/2019. S. Solanki acknowledges the Council of Scientific and Industrial Research (CSIR), India for the award of a research associate fellowship.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kucherenko, I.; Soldatkin, O.; Kucherenko, D.Y.; Soldatkina, O.; Dzyadevych, S.V. Advances in nanomaterial application in enzyme-based electrochemical biosensors: A review. *Nanoscale Adv.* **2019**, *1*, 4560–4577.
2. Pandey, P.; Datta, M.; Malhotra, B. Prospects of nanomaterials in biosensors. *Anal. Lett.* **2008**, *41*, 159–209.
3. Tomczykowa, M.; Plonska-Brzezinska, M.E. Conducting polymers, hydrogels and their composites: Preparation, properties and bioapplications. *Polymers* **2019**, *11*, 350.

4. Heinze, J.; Frontana-Uribe, B.A.; Ludwigs, S. Electrochemistry of conducting polymers--persistent models and new concepts. *Chem. Rev.* **2010**, *110*, 4724–4771.
5. Kankare, J. Electronically conducting polymers: Basic methods of synthesis and characterization. *Plast. Eng. N. Y.* **1998**, *45*, 167–200.
6. Foulds, N.C.; Lowe, C.R. Enzyme entrapment in electrically conducting polymers. Immobilisation of glucose oxidase in polypyrrole and its application in amperometric glucose sensors. *J. Chem. Soc. Faraday Trans. 1 Phys. Chem. Condens. Phases* **1986**, *82*, 1259–1264.
7. Umana, M.; Waller, J. Protein-modified electrodes. The glucose oxidase/polypyrrole system. *Anal. Chem.* **1986**, *58*, 2979–2983.
8. Bartlett, P.; Whitaker, R. Electrochemical immobilisation of enzymes: Part II. Glucose oxidase immobilised in poly-N-methylpyrrole. *J. Eectroanal. Chem. Inter. Electrochem.* **1987**, *224*, 37–48.
9. Dong, H.; Li, C.M.; Chen, W.; Zhou, Q.; Zeng, Z.X.; Luong, J.H.T. Sensitive amperometric immunosensing using polypyrrolepropylic acid films for biomolecule immobilization. *Anal. Chem.* **2006**, *78*, 7424–7431.
10. Oztekin, Y.; Ramanaviciene, A.; Yazicigil, Z.; Solak, A.O.; Ramanavicius, A. Direct electron transfer from glucose oxidase immobilized on polyphenanthroline-modified glassy carbon electrode. *Biosens. Bioelectron.* **2011**, *26*, 2541–2546.
11. Kausaite-Minkstimiene, A.; Mazeiko, V.; Ramanaviciene, A.; Ramanavicius, A. Evaluation of amperometric glucose biosensors based on glucose oxidase encapsulated within enzymatically synthesized polyaniline and polypyrrole. *Sens. Act. B Chem.* **2011**, *158*, 278–285.
12. Riddick, J.; Bunger, W.; Sakano, T. Organic Solvents. In *Techniques of Chemistry*, 4th ed.; John Wiley Sons: Hoboken, NJ, USA, 1985; Volume II.
13. Sax, N.I.; Hawley, G.G.; Lewis, R.J. *Hawley's Condensed Chemical Dictionary*; Chapman-Hall: Atlanta, GA, USA, 1987.
14. Diaz, A.F.; Castillo, J.I.; Logan, J.; Lee, W.-Y. Electrochemistry of conducting polypyrrole films. *J. Eectroanal. Chem. Inter. Electrochem.* **1981**, *129*, 115–132.
15. Paramo, U.; Ibanez, J.G. Electrochemical modulation of the thickness of polypyrrole films by using different anionic dopants. *Int. J. Electrochem. Sci.* **2011**, *6*, 5172–5188.
16. Silk, T.; Hong, Q.; Tamm, J.; Compton, R.G. AFM studies of polypyrrole film surface morphology I. The influence of film thickness and dopant nature. *Synth. Met.* **1998**, *93*, 59–64.
17. Satoh, M.; Kaneto, K.; Yoshino, K. Dependences of electrical and mechanical properties of conducting polypyrrole films on conditions of electrochemical polymerization in an aqueous medium. *Synth. Met.* **1986**, *14*, 289–296.
18. Zhou, M.; Heinze, J. Electropolymerization of pyrrole and electrochemical study of polypyrrole. 3. Nature of "water effect" in acetonitrile. *J. Phys. Chem. B* **1999**, *103*, 8451–8457.
19. Williams, A.; Ibrahim, I.T. Carbodiimide chemistry: Recent advances. *Chem. Rev.* **1981**, *81*, 589–636.
20. McLeod, G.G.; Mahboubian-Jones, M.; Pethrick, R.A.; Watson, S.D.; Truong, N.; Galin, J.; Francois, J. Synthesis, electrochemical polymerization and properties of poly (2,5-di-(-2-thienyl)-pyrrole). *Polymer* **1986**, *27*, 455–458.
21. Kim, S.; Jang, L.K.; Park, H.S.; Lee, J.Y. Electrochemical deposition of conductive and adhesive polypyrrole-dopamine films. *Sci. Rep.* **2016**, *6*, 30475.
22. Ito, S.; Murata, K.; Teshima, S.; Aizawa, R.; Asako, Y.; Takahashi, K.; Hoffman, B.M. Simple synthesis of water-soluble conducting polyaniline. *Synth. Met.* **1998**, *96*, 161–163.
23. Bhadra, S.; Khastgir, D.; Singha, N.K.; Lee, J.H. Progress in preparation, processing and applications of polyaniline. *Prog. Pol. Sci.* **2009**, *34*, 783–810.
24. Nagarajan, R.; Tripathy, S.; Kumar, J.; Bruno, F.F.; Samuelson, L. An enzymatically synthesized conducting molecular complex of polyaniline and poly (vinylphosphonic acid). *Macromolecules* **2000**, *33*, 9542–9547.
25. Wallace, G.G.; Teasdale, P.R.; Spinks, G.M.; Kane-Maguire, L.A. *Conductive Electroactive Polymers: Intelligent Materials Systems*; CRC Press: Boca Raton, FL, USA, 2002.
26. Zotti, G.; Cattarin, S.; Comisso, N. Electrodeposition of polythiophene, polypyrrole and polyaniline by the cyclic potential sweep method. *J. Eectroanal. Chem. Inter. Electrochem.* **1987**, *235*, 259–273.
27. Mu, S.; Chen, C.; Wang, J. The kinetic behavior for the electrochemical polymerization of aniline in aqueous solution. *Synth. Met.* **1997**, *88*, 249–254.
28. Park, S.-M.; Lee, H.-J. Recent advances in electrochemical studies of π-conjugated polymers. *Bull. Korean Chem. Soc.* **2005**, *26*, 697–706.

29. Karyakin, A.; Strakhova, A.; Yatsimirsky, A. Self-doped polyanilines electrochemically active in neutral and basic aqueous solutions: Electropolymerization of substituted anilines. *J. Electroanal. Chem.* **1994**, *371*, 259–265.
30. Mu, S. Synthesis of poly (aniline-co-5-aminosalicylic acid) and its properties. *Synth. Met.* **2011**, *161*, 1306–1312.
31. Obaid, A.; El-Mossalamy, E.; Al-Thabaiti, S.; El-Hallag, I.; Hermas, A.; Asiri, A. Electrodeposition and characterization of polyaniline on stainless steel surface via cyclic, convolutive voltammetry and SEM in aqueous acidic solutions. *Int. J. Electrochem. Sci.* **2014**, *9*, 1003–1015.
32. Kobayashi, T.; Yoneyama, H.; Tamura, H. Electrochemical reactions concerned with electrochromism of polyaniline film-coated electrodes. *J. Eectroanal. Chem. Inter. Electrochem.* **1984**, *177*, 281–291.
33. Cesarino, I.; Moraes, F.C.; Machado, S.A. A biosensor based on polyaniline-carbon nanotube core-shell for electrochemical detection of pesticides. *Electroanalysis* **2011**, *23*, 2586–2593.
34. Shoaie, N.; Daneshpour, M.; Azimzadeh, M.; Mahshid, S.; Khoshfetrat, S.M.; Jahanpeyma, F.; Gholaminejad, A.; Omidfar, K.; Foruzandeh, M. Electrochemical sensors and biosensors based on the use of polyaniline and its nanocomposites: A review on recent advances. *Microchim. Acta* **2019**, *186*, 465.
35. Tourillon, G.; Garnier, F. Morphology of conducting organic polymers: Polythiophene and poly (3-methyl thiophene). *J. Polym. Sci. Polym. Phys. Ed.* **1984**, *22*, 33–39.
36. Aguilar-Ortíz, E.; Zaragoza-Galán, G.; Solladié, N.; Rein, R.; Aguilar-Martínez, M.; Macías-Ruvalcaba, N.; Rivera, E. Preparation and characterization of novel polythiophenes bearing oligo (ethylene glycol) spacers and porphyrin units: Optical and electrochemical properties. *Synth. Met.* **2012**, *162*, 1000–1009.
37. Wang, F.; Li, M.; Wang, B.; Zhang, J.; Cheng, Y.; Liu, L.; Lv, F.; Wang, S. Synthesis and characterization of water-soluble polythiophene derivatives for cell imaging. *Sci. Rep.* **2015**, *5*, 7617.
38. Heller, A.; Feldman, B. Electrochemical glucose sensors and their applications in diabetes management. *Chem. Rev*. **2008**, *108*, 2482–2505.
39. Elgrishi, N.; Rountree, K.J.; McCarthy, B.B.; Rountree, E.S.; Eisenhart, T.T.; Dempsey, J.L. A practical beginner's guide to cyclic voltammetry. *J. Chem. Educ*. **2018**, *95*, 197–206.
40. Romani, A.; Minunni, M.; Mulinacci, N.; Pinelli, P.; Vincieri, F.F.; Del Carlo, M.; Mascini, M. Comparison among differential pulse voltammetry, amperometric biosensor, and HPLC/DAD analysis for polyphenol determination. *J. Agric. Food Chem*. **2000**, *48*, 1197–1203.
41. Velychko, T.P.; Soldatkin, O.O.; Melnyk, V.G.; Marchenko, S.V.; Kirdeciler, S.K.; Akata, B.; Soldatkin, A.P.; Elskaya, A.V.; Dzyadevych, S.V. A novel conductometric urea biosensor with improved analytical characteristic based on recombinant urease adsorbed on nanoparticle of silicalite. *Nanoscale Res. Lett*. **2016**, *11*, 106.
42. Xiao, C.; Lachance, B.; Sunahara, G.; Luong, J.H.T. Assessment of cytotoxicity using electric cell− substrate impedance sensing: Concentration and time response function approach. *Anal. Chem*. **2002**, *74*, 5748–5753.
43. Male, K.B.; Lachance, B.; Hrapovic, S.; Sunahara, G.; Luong, J.H.T. Assessment of cytotoxicity of quantum dots and gold nanoparticles using cell-based impedance spectroscopy. *Anal. Chem*. **2018**, *80*, 5487–5493.
44. Allen, J.B.; Faulkner, L.R. *Electrochemical Methods Fundamentals and Applications*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2001.
45. Ronkainen, N.J.; Halsall, H.B.; Heineman, W.R. Electrochemical biosensors. *Chem. Soc. Rev*. **2010**, *39*, 1747–1763.
46. Scouten, W.H.; Luong, J.H.T.; Brown, R.S. Enzyme or protein immobilization techniques for applications in biosensor design. *Trends Biotechnol*. **1995**, *13*, 178–185.
47. Welch, M.E.; Doublet, T.; Bernard, C.; Malliaras, G.G.; Ober, C.K. A glucose sensor via stable immobilization of the GOx enzyme on an organic transistor using a polymer brush. *Poly. Chem. Sci. A-Poly. Chem*. **2015**, *53*, 372–377.
48. Gu, H.; Xiao, X.D.; Loh, K.P. Electrochemical impedance sensing of DNA hybridization on conducting polymer film-modified diamond. *J. Phys. Chem. B* **2005**, *109*, 13611–13618.
49. Setti, L.; Fraleoni-Morgera, A.; Mencarelli, I.; Filippini, A.; Ballarin, B.; Di Biase, M. An HRP-based amperometric biosensor fabricated by thermal inkjet printing. *Sens. Actuators B Chem.* **2007**, *126*, 252–257.
50. Safavi, A.; Maleki, N.; Farjami, E. Fabrication of a glucose sensor based on a novel nanocomposite electrode. *Biosens. Bioelectron*. **2009**, *24*, 1655–1660.
51. Derkus, B. Applying the miniaturization technologies for biosensor design. *Biosens. Bioeletron*. **2016**, *79*, 736–741.

52. Wilson, G.S.; Gifford, R. Biosensors for real-time in vivo measurements. *Biosens. Bioelectron.* **2005**, *20*, 2388–2403.
53. Mello, H.; Mulato, M. Enzymatically functionalized polyaniline thin films produced with one-step electrochemical immobilization and its application in glucose and urea potentiometric biosensors. *Biomed. Microdevices* **2020**, *22*, 22.
54. Erol, E.; Yildirim, E.; Cete, S. Construction of biosensor for hypoxanthine determination by immobilization of xanthine oxidase and uricase in polypyrrole-paratoluenesulfonate film. *J. Solid State Electrochem.* **2020**, *24*, 1695–1707.
55. Nessark, F.; Eissa, M.; Baraket, A.; Zine, N.; Nessark, B.; Zouaoui, A.; Bausells, J.; Errachid, A. Capacitance polypyrrole-based impedimetric immunosensor for interleukin-10 cytokine detection. *Electroanalysis* **2020**, *32*, 1795–1806.
56. Kumar, S.; Umar, M.; Saifi, A.; Kumar, S.; Augustine, S.; Srivastava, S.; Malhotra, B.D. Electrochemical paper based cancer biosensor using iron oxide nanoparticles decorated PEDOT: PSS. *Anal. Chim. Acta* **2019**, *1056*, 135–145.
57. Kumar, S.; Kumar, S.; Srivastava, S.; Yadav, B.K.; Lee, S.H.; Sharma, J.G.; Doval, D.C.; Malhotra, B.D. Reduced graphene oxide modified smart conducting paper for cancer biosensor. *Biosens. Bioelectron.* **2015**, *73*, 114–122.
58. Spychalska, K.; Zając, D.; Cabaj, J. Electrochemical biosensor for detection of 17β-estradiol using semi-conducting polymer and horseradish peroxidase. *RSC Adv.* **2020**, *10*, 9079–9087.
59. Baluta, S.; Świst, A.; Cabaj, J.; Malecha, K. Point-of-care testing–biosensor for norepinephrine determination. *Int. J. Electron. Telecommun.* **2020**, *66*, 369–372.
60. Goda, T.; Miyahara, Y. Electrodeposition of Zwitterionic PEDOT films for conducting and antifouling surfaces. *Langmuir* **2019**, *35*, 1126–1133.
61. Wang, J.; Wang, D.; Hui, N. A low fouling electrochemical biosensor based on the zwitterionic polypeptide doped conducting polymer PEDOT for breast cancer marker BRCA1 detection. *Bioelectrochemistry* **2020**, *136*, 107595.
62. Dabrowski, M.; Lach, P.; Cieplak, M.; Kutner, W. Nanostructured molecularly imprinted polymers for protein chemosensing. *Biosens. Bioelectron.* **2018**, *102*, 17–26.
63. Ayankojo, A.G.; Reut, J.; Opik, A.; Furchner, A.; Syritski, V. Hybrid molecularly imprinted polymer for amoxicillin detection. *Biosens. Bioelectron.* **2018**, *118*, 102–107.
64. Waffo, A.; Yesildag, C.; Caserta, G.; Katz, S.; Zebger, I.; Lensen, M.; Wollenberger, U.; Scheller, F.; Altintas, Z. Fully electrochemical MIP sensor for artemisinin. *Sens. Actuators B Chem.* **2018**, *275*, 163–173.
65. Kane-Maguire, L.A.; Wallace, G.G. Chiral conducting polymers. *Chem. Soc. Rev.* **2010**, *39*, 2545–2576.
66. Shang, X.; Song, I.; Jung, G.Y.; Choi, W.; Ohtsu, H.; Lee, J.H.; Koo, J.Y.; Liu, B.; Ahn, J.; Kawano, M. Chiral self-sorted multifunctional supramolecular biocoordination polymers and their applications in sensors. *Nat. Commun.* **2018**, *9*, 1–12.
67. Lee, K.M.; Kim, K.H.; Yoon, H.; Kim, H. Chemical design of functional polymer structures for biosensors: From nanoscale to macroscale. *Polymers* **2018**, *10*, 551.
68. Naficy, S.; Oveissi, F.; Patrick, B.; Schindeler, A.; Dehghani, F. Printed, flexible pH sensor hydrogels for wet environments. *Adv. Mat. Technol.* **2018**, *3*, 1800137.
69. Culver, H.R.; Wechsler, M.E.; Peppas, N.A. Label-free detection of tear biomarkers using hydrogel-coated gold nanoshells in a localized surface plasmon resonance-based biosensor. *ACS Nano* **2018**, *12*, 9342–9354.
70. Li, L.; Pan, L.; Ma, Z.; Yan, K.; Cheng, W.; Shi, Y.; Yu, G. All inkjet-printed amperometric multiplexed biosensors based on nanostructured conductive hydrogel electrodes. *Nano Lett.* **2018**, *18*, 3322–3327.
71. Meng, L.; Turner, A.P.; Mak, W.C. Tunable 3D nanofibrous and bio-functionalised PEDOT network explored as a conducting polymer-based biosensor. *Biosens. Bioelectron.* **2020**, *159*, 112181.
72. Słoniewska, A.; Kasztelan, M.; Berbeć, S.; Pałys, B. Influence of buffer solution on structure and electrochemical properties of poly (3, 4-ethylenedioxythiophene)/poly (styrenesulfonate) hydrogels. *Synth. Met.* **2020**, *263*, 116363.
73. Cete, S.; Ozyurt, M.; Yildirim, E.; Akin, D. A novel biosensor with the use of polypyrrole–poly (sodium-4-styrenesulphonate) as a dopant in the determination of glucose. *Chem. Papers* **2020**, *74*, 799–808.
74. Kajisa, T.; Li, W.; Michinobu, T.; Sakata, T. Well-designed dopamine-imprinted polymer interface for selective and quantitative dopamine detection among catecholamines using a potentiometric biosensor. *Biosens. Bioelectron.* **2018**, *117*, 810–817.

75. Ashley, J.; Feng, X.; Sun, Y. A multifunctional molecularly imprinted polymer-based biosensor for direct detection of doxycycline in food samples. *Talanta* **2018**, *182*, 49–54.
76. Ribeiro, J.; Pereira, C.; Silva, A.; Sales, M.G.F. Disposable electrochemical detection of breast cancer tumour marker CA 15-3 using poly (toluidine blue) as imprinted polymer receptor. *Biosens. Bioelectron.* **2018**, *109*, 246–254.
77. Palladino, P.; Minunni, M.; Scarano, S. Cardiac Troponin T capture and detection in real-time via epitope-imprinted polymer and optical biosensing. *Biosens. Bioelectron.* **2018**, *106*, 93–98.
78. Roushani, M.; Nezhadali, A.; Jalilian, Z. An electrochemical chlorpyrifos aptasensor based on the use of a glassy carbon electrode modified with an electropolymerized aptamer-imprinted polymer and gold nanorods. *Microchim. Acta* **2018**, *185*, 551.
79. Liu, S.; Ma, Y.; Cui, M.; Luo, X. Enhanced electrochemical biosensing of alpha-fetoprotein based on three-dimensional macroporous conducting polymer polyaniline. *Sens. Actuators B Chem.* **2018**, *255*, 2568–2574.
80. Lin, Z.T.; Li, Y.; Gu, J.; Wang, H.; Zhu, Z.; Hong, X.; Zhang, Z.; Lu, Q.; Qiu, J.; Wang, X. A conductive nanowire-mesh biosensor for ultrasensitive detection of serum c-reactive protein in melanoma. *Adv. Func. Mat.* **2018**, *28*, 1802482.
81. Wei, Y.; Zeng, Q.; Hu, Q.; Wang, M.; Tao, J.; Wang, L. Self-cleaned electrochemical protein imprinting biosensor basing on a thermo-responsive memory hydrogel. *Biosens. Bioelectron.* **2018**, *99*, 136–141.
82. Ekomo, V.M.; Branger, C.; Bikanga, R.; Florea, A.M.; Istamboulie, G.; Calas-Blanchard, C.; Noguer, T.; Sarbu, A.; Brisset, H. Detection of Bisphenol A in aqueous medium by screen printed carbon electrodes incorporating electrochemical molecularly imprinted polymers. *Biosens. Bioelectron.* **2018**, *112*, 156–161.
83. Meng, L.; Turner, A.P.F.; Mak, W.C. Positively-charged hierarchical PEDOT interface with enhanced electrode kinetics for NADH-based biosensors. *Biosens. Bioelectron.* **2018**, *120*, 115–121.
84. Kuralay, F.; Dukar, N.; Bayramli, Y. Poly-L-lysine coated surfaces for ultrasensitive nucleic acid detection. *Electroanalysis* **2018**, *30*, 1556–1565.
85. Moreira, F.T.; Truta, L.A.; Sales, M.G.F. Biomimetic materials assembled on a photovoltaic cell as a novel biosensing approach to cancer biomarker detection. *Sci. Rep.* **2018**, *8*, 1–11.
86. Kausaite-Minkstimiene, A.; Glumbokaite, L.; Ramanaviciene, A.; Dauksaite, E.; Ramanavicius, A. An amperometric glucose biosensor based on poly (pyrrole-2-carboxylic acid)/glucose oxidase biocomposite. *Electroanalysis* **2018**, *30*, 1642–1652.
87. Wu, H.; Lee, C.J.; Wang, H.; Hu, Y.; Young, M.; Han, Y.; Xu, F.J.; Cong, H.; Cheng, G. Highly sensitive and stable zwitterionic poly (sulfobetaine-3,4-ethylenedioxythiophene) (PSBEDOT) glucose biosensor. *Chem. Sci.* **2018**, *9*, 2540–2546.
88. Soganci, T.; Soyleyici, H.C.; Demirkol, D.O.; Ak, M.; Timur, S. Use of super-structural conducting polymer as functional immobilization matrix in biosensor design. *J. Electrochem. Soc.* **2018**, *165*, B22.
89. Apetrei, R.-M.; Cârâc, G.; Bahrim, G.; Camurlu, P. Sensitivity enhancement for microbial biosensors through cell Self-Coating with polypyrrole. *Int. J. Polym. Mat. Polym. Biomat.* **2018**, *68*, 1058–1067.
90. Apetrei, R.-M.; Cârâc, G.; Bahrim, G.; Camurlu, P. Utilization of enzyme extract self-encapsulated within polypyrrole in sensitive detection of catechol. *Enzym. Microb. Technol.* **2019**, *128*, 34–39.
91. Jakhar, P.; Shukla, M.; Singh, V. Electrochemically synthesized polypyrrole nanofibers towards glucose biosensing application. In Proceedings of the International Workshop on the Physics of Semiconductor and Devices, Delhi, India, 12–15 December 2017; 215, pp. 871–874.
92. Soylemez, S.; Bekmezci, S.A.; Goker, S.; Toppare, L. Fabrication of a novel polymeric scaffold for amperometric laccase biosensor. *J. Polym. Sci. Part A Polym. Chem.* **2019**, *57*, 2333–2339.
93. Li, J.; Bi, X.; Tamulevičius, S.; Erts, D.; Chang, C.-F.; Gu, Y. Fabrication of a biocompatible and continuous glucose biosensor with the poly (3, 4-ethylenedioxythiophene) modified electrode. *J. Taiwan Inst. Chem. Eng.* **2019**, *104*, 1–7.
94. Cevik, E.; Cerit, A.; Tombuloglu, H.; Sabit, H.; Yildiz, H.B. Electrochemical glucose biosensors: Whole cell microbial and enzymatic determination based on 10-(4H-dithieno [3, 2-b: 2′, 3′-d] pyrrol-4-yl)decan-1-amine interfaced glassy carbon electrodes. *Anal. Lett.* **2019**, *52*, 1138–1152.
95. Ranjbar, S.; Nejad, M.A.F.; Parolo, C.; Shahrokhian, S.; Merkoçi, A. Smart chip for visual detection of bacteria using the electrochromic properties of polyaniline. *Anal. Chem.* **2019**, *91*, 14960–14966.
96. Hosseinian, M.; Najafpour, G.; Rahimpour, A. Amperometric urea biosensor based on immobilized urease on polypyrrole and macroporous polypyrrole modified Pt electrode. *Turk. J. Chem.* **2019**, *43*, 1063–1074.

97. Fenoy, G.E.; Marmisolle, W.A.; Azzaroni, O.; Knoll, W. Acetylcholine biosensor based on the electrochemical functionalization of graphene field-effect transistors. *Biosens. Bioelectron.* **2020**, *148*, 111796.
98. Gülercan, D.; Gergin, İ.; Sarac, A.S. Preparation and Electrochemical Performances of graphene oxide/PEDOT and reduced graphene oxide/PEDOT nanofibers and nanocomposites. *Fibers Polym.* **2018**, *19*, 2178–2187.
99. Hussain, K.K.; Gurudatt, N.; Akhtar, M.H.; Seo, K.-D.; Park, D.-S.; Shim, Y.-B. Nano-biosensor for the in vitro lactate detection using bi-functionalized conducting polymer/N, S-doped carbon; the effect of αCHC inhibitor on lactate level in cancer cell lines. *Biosens. Bioelectron.* **2020**, *155*, 112094.
100. Seo, K.-D.; Hossain, M.M.; Gurudatt, N.; Choi, C.S.; Shiddiky, M.J.; Park, D.-S.; Shim, Y.-B. Microfluidic neurotransmitters sensor in blood plasma with mediator-immobilized conducting polymer/N, S-doped porous carbon composite. *Sens. Actuators B Chem.* **2020**, *313*, 128017.
101. Ma, L.; Zhou, L.; He, Y.; Wang, L.; Huang, Z.; Jiang, Y.; Gao, J. Hierarchical nanocomposites with an N-doped carbon shell and bimetal core: Novel enzyme nanocarriers for electrochemical pesticide detection. *Biosens. Bioelectron.* **2018**, *121*, 166–173.
102. Zheng, W.; Wu, H.; Jiang, Y.; Xu, J.; Li, X.; Zhang, W.; Qiu, F. A molecularly-imprinted-electrochemical-sensor modified with nano-carbon-dots with high sensitivity and selectivity for rapid determination of glucose. *Anal. Biochem.* **2018**, *555*, 42–49.
103. Zhou, Z.; Zhang, Y.; Shen, Y.; Liu, S.; Zhang, Y. Molecular engineering of polymeric carbon nitride: Advancing applications from photocatalysis to biosensing and more. *Chem. Soc. Rev.* **2018**, *47*, 2298–2321.
104. Yan, K.; Yang, Y.; Zhang, J. A self-powered sensor based on molecularly imprinted polymer-coupled graphitic carbon nitride photoanode for selective detection of bisphenol A. *Sens. Actuators B Chem.* **2018**, *259*, 394–401.
105. Han, R.; Wang, G.; Xu, Z.; Zhang, L.; Li, Q.; Han, Y.; Luo, X. Designed antifouling peptides planted in conducting polymers through controlled partial doping for electrochemical detection of biomarkers in human serum. *Biosens. Bioelectron.* **2020**, *164*, 112317.
106. Yilmaz, Y.Y.; Yalcinkaya, E.E.; Demirkol, D.O.; Timur, S. 4-aminothiophenol-intercalated montmorillonite: Organic-inorganic hybrid material as an immobilization support for biosensors. *Sens. Actuators B Chem.* **2020**, *307*, 127665.
107. Phongphut, A.; Chayasombat, B.; Cass, A.E.; Sirisuk, A.; Phisalaphong, M.; Prichanont, S.; Thanachayanont, C. Clay/au nanoparticle composites as acetylcholinesterase carriers and modified-electrode materials: A comparative study. *Appl. Clay Sci.* **2020**, *194*, 105704.
108. Zheng, H.; Liu, M.; Yan, Z.; Chen, J. Highly selective and stable glucose biosensor based on incorporation of platinum nanoparticles into polyaniline-montmorillonite hybrid composites. *Microchem. J.* **2020**, *152*, 104266.
109. Erkmen, C.; Kurbanoglu, S.; Uslu, B. Fabrication of poly (3, 4-ethylenedioxythiophene)-iridium oxide nanocomposite based Tyrosinase biosensor for the dual detection of catechol and azinphos methyl. *Sens. Actuators B Chem.* **2020**, *316*, 128121.
110. Altun, A.; Apetreia, R.M.; Camurlu, P. Reagentless amperometric glucose biosensors: Ferrocene-tethering and copolymerization. *J. Electrochem. Soc.* **2020**, *167*, 107507.
111. Sun, Y.; Zeng, W.; Sun, H.; Luo, S.; Chen, D.; Chan, V.; Liao, K. Inorganic/polymer-graphene hybrid gel as versatile electrochemical platform for electrochemical capacitor and biosensor. *Carbon* **2018**, *132*, 589–597.
112. Wang, X.; Zhu, Y.; Olsen, T.R.; Sun, N.; Zhang, W.; Pei, R.; Lin, Q. A graphene aptasensor for biomarker detection in human serum. *Electrochim. Acta* **2018**, *290*, 356–363.
113. Fusco, G.; Gobel, G.; Zanoni, R.; Bracciale, M.P.; Favero, G.; Mazzei, F.; Lisdat, F. Aqueous polythiophene electrosynthesis: A new route to an efficient electrode coupling of PQQ-dependent glucose dehydrogenase for sensing and bioenergetic applications. *Biosens. Bioelectron.* **2018**, *112*, 8–17.
114. Gautam, V.; Singh, K.P.; Yadav, V.L. Polyaniline/MWCNTs/starch modified carbon paste electrode for non-enzymatic detection of cholesterol: Application to real sample (cow milk). *Anal. Bioanal. Chem.* **2018**, *410*, 2173–2181.
115. Liu, Y.; Turner, A.P.F.; Zhao, M.; Mak, W.C. Processable enzyme-hybrid conductive polymer composites for electrochemical biosensing. *Biosens. Bioelectron.* **2018**, *100*, 374–381.
116. Nguyet, N.T.; Van Thu, V.; Lan, H.; Trung, T.; Le, A.-T.; Pham, V.-H.; Tam, P.D. Simple label-free dna sensor based on CeO_2 nanorods decorated with ppy nanoparticles. *J. Electron. Mat.* **2019**, *48*, 6231–6239.

117. Yang, T.; Ren, X.; Yang, M.; Li, X.; He, K.; Rao, A.; Wan, Y.; Yang, H.; Wang, S.; Luo, Z. A highly sensitive label-free electrochemical immunosensor based on poly (indole-5-carboxylicacid) with ultra-high redox stability. *Biosens. Bioelectron.* **2019**, *141*, 111406.
118. Thunyakontirakun, W.; Sriwichai, S.; Phanichphant, S.; Janmanee, R. Fabrication of poly (pyrrole-3-carboxylic acid)/graphene oxide composite thin film for glucose biosensor. *Mat. Today Proc.* **2019**, *17*, 2070–2077.
119. Mousa, H.M.; Aggas, J.R.; Guiseppi-Elie, A. Electropolymerization of aniline and (N-phenyl-o-phenylenediamine) for glucose biosensor application. *Mat. Lett.* **2019**, *238*, 267–270.
120. Salazar, P.; Martín, M.; González-Mora, J.L. In situ electrodeposition of cholesterol oxidase-modified polydopamine thin film on nanostructured screen printed electrodes for free cholesterol determination. *J. Electroanal. Chem.* **2019**, *837*, 191–199.
121. Wang, J.; Hui, N. Electrochemical functionalization of polypyrrole nanowires for the development of ultrasensitive biosensors for detecting microRNA. *Sens. Actuators B Chem.* **2019**, *281*, 478–485.
122. Abd-Wahab, F.; Abdul Guthoos, H.F.; Wan Salim, W.W.A. Solid-state rGO-PEDOT: PSS transducing material for cost-effective enzymatic sensing. *Biosensors* **2019**, *9*, 36.
123. Sahyar, B.Y.; Kaplan, M.; Ozsoz, M.; Celik, E.; Otles, S. Electrochemical xanthine detection by enzymatic method based on Ag doped ZnO nanoparticles by using polypyrrole. *Bioelectrochemistry* **2019**, *130*, 107327.
124. Altun, A.; Apetrei, R.M.; Camurlu, P. The effect of copolymerization and carbon nanoelements on the performance of poly (2,5-di(thienyl)pyrrole) biosensors. *Mat. Sci. Eng. C Mater. Biol. Appl.* **2019**, *105*, 110069.
125. Yashas, S.R.; Sandeep, S.; Shivakumar, B.P.; Swamy, N.K. A matrix of perovskite micro-seeds and polypyrrole nanotubes tethered laccase/graphite biosensor for sensitive quantification of 2, 4-dichlorophenol in wastewater. *Anal. Methods* **2019**, *11*, 4511–4519.
126. Ahmad, L.; Salmon, L.; Korri-Youssoufi, H. Electrochemical detection of the human cancer biomarker 'autocrine motility factor-phosphoglucose isomerase' based on a biosensor formed with a monosaccharidic inhibitor. *Sens. Actuators B Chem.* **2019**, *299*, 126933.
127. Aghamiri, Z.S.; Mohsennia, M.; Rafiee-Pour, H.-A. Immobilization of cytochrome C on polyaniline/polypyrrole/carboxylated multi-walled carbon nanotube/glassy carbon electrode: Biosensor fabrication. *J. Solid State Electrochem.* **2019**, *23*, 2233–2242.
128. Almeida, L.C.; Correia, R.D.; Marta, A.; Squillaci, G.; Morana, A.; La Cara, F.; Correia, J.P.; Viana, A.S. Electrosynthesis of polydopamine films-tailored matrices for laccase-based biosensors. *Appl. Surf. Sci.* **2019**, *480*, 979–989.
129. Phetsang, S.; Jakmunee, J.; Mungkornasawakul, P.; Laocharoensuk, R.; Ounnunkad, K. Sensitive amperometric biosensors for detection of glucose and cholesterol using a platinum/reduced graphene oxide/poly (3-aminobenzoic acid) film-modified screen-printed carbon electrode. *Bioelectrochemistry* **2019**, *127*, 125–135.
130. Nguyen, T.N.; Nolan, J.K.; Park, H.; Lam, S.; Fattah, M.; Page, J.C.; Joe, H.-E.; Jun, M.B.; Lee, H.; Kim, S.J. Facile fabrication of flexible glutamate biosensor using direct writing of platinum nanoparticle-based nanocomposite ink. *Biosens. Bioelectron.* **2019**, *131*, 257–266.
131. Gupta, A.; Bhardwaj, S.K.; Sharma, A.L.; Kim, K.-H.; Deep, A. Development of an advanced electrochemical biosensing platform for *E. coli* using hybrid metal-organic framework/polyaniline composite. *Environ. Res.* **2019**, *171*, 395–402.
132. Lee, W.C.; Kim, K.B.; Gurudatt, N.G.; Hussain, K.K.; Choi, C.S.; Park, D.S.; Shim, Y.B. Comparison of enzymatic and non-enzymatic glucose sensors based on hierarchical Au-Ni alloy with conductive polymer. *Biosens. Bioelectron.* **2019**, *130*, 48–54.
133. Almeida, L.; Correia, R.; Squillaci, G.; Morana, A.; La Cara, F.; Correia, J.; Viana, A. Electrochemical deposition of bio-inspired laccase-polydopamine films for phenolic sensors. *Electrochim. Acta* **2019**, *319*, 462–471.
134. Maity, D.; Minitha, C.R.; Rajendra Kumar, R.T. Glucose oxidase immobilized amine terminated multiwall carbon nanotubes/reduced graphene oxide/polyaniline/gold nanoparticles modified screen-printed carbon electrode for highly sensitive amperometric glucose detection. *Mat. Sci. Eng. C-Mat. Biol. Appl.* **2019**, *105*, 110075.
135. Keteklahijani, Y.Z.; Sharif, F.; Roberts, E.P.; Sundararaj, U. Enhanced sensitivity of dopamine biosensors: An electrochemical approach based on nanocomposite electrodes comprising polyaniline, nitrogen-doped graphene, and DNA-functionalized carbon nanotubes. *J. Electrochem. Soc.* **2019**, *166*, B1415.

136. Li, Z.; Yin, J.; Gao, C.; Qiu, G.; Meng, A.; Li, Q. The construction of electrochemical aptasensor based on coral-like poly-aniline and Au nano-particles for the sensitive detection of prostate specific antigen. *Sens. Actuators B Chem.* **2019**, *295*, 93–100.
137. Al-Sagur, H.; Shanmuga Sundaram, K.; Kaya, E.N.; Durmus, M.; Basova, T.V.; Hassan, A. Amperometric glucose biosensing performance of a novel graphene nanoplatelets-iron phthalocyanine incorporated conducting hydrogel. *Biosens. Bioelectron.* **2019**, *139*, 111323.
138. Alim, S.; Kafi, A.K.M.; Rajan, J.; Yusoff, M.M. Application of polymerized multiporous nanofiber of SnO_2 for designing a bienzyme glucose biosensor based on HRP/GOx. *Int. J. Biol. Macromol.* **2019**, *123*, 1028–1034.
139. Guo, X.; Li, Y.; Li, Y.; Ye, Z.; Zhang, J.; Zhu, T.; Li, F. An L012@ PAni-PAAm hydrogel composite based-electrochemiluminescence biosensor for in situ detection of H_2O_2 released from cardiomyocytes. *Electrochim. Acta* **2020**, *354*, 136763.
140. Khoder, R.; Korri-Youssoufi, H. E-DNA biosensors of M. tuberculosis based on nanostructured polypyrrole. *Mat. Sci. Eng. C-Mat. Biol. Appl.* **2020**, *108*, 110371.
141. Aydin, M.; Aydin, E.B.; Sezginturk, M.K. A highly selective electrochemical immunosensor based on conductive carbon black and star PGMA polymer composite material for IL-8 biomarker detection in human serum and saliva. *Biosens. Bioelectron.* **2018**, *117*, 720–728.
142. He, L.; Cui, B.; Liu, J.; Song, Y.; Wang, M.; Peng, D.; Zhang, Z. Novel electrochemical biosensor based on core-shell nanostructured composite of hollow carbon spheres and polyaniline for sensitively detecting malathion. *Sens. Actuators B Chem.* **2018**, *258*, 813–821.
143. Tancharoen, C.; Sukjee, W.; Thepparit, C.; Jaimipuk, T.; Auewarakul, P.; Thitithanyanont, A.; Sangma, C. Electrochemical biosensor based on surface imprinting for zika virus detection in serum. *ACS Sens.* **2018**, *4*, 69–75.
144. Yang, J.; Liu, Y.; Liu, S.; Li, L.; Zhang, C.; Liu, T. Conducting polymer composites: Material synthesis and applications in electrochemical capacitive energy storage. *Mat. Chem. Front.* **2017**, *1*, 251–268.
145. Liu, Y.; Lu, N.; Poyraz, S.; Wang, X.; Yu, Y.; Scott, J.; Smith, J.; Kim, M.J.; Zhang, X. One-pot formation of multifunctional Pt-conducting polymer intercalated nanostructures. *Nanoscale* **2013**, *5*, 3872–3879.
146. Shahrokhian, S.; Shamloofard, M.; Salimian, R. Investigation of the electrochemical behavior of tizanidine on the surface of glassy carbon electrode modified with multi-walled carbon nanotube/titan yellow–doped polypyrrole. *J. Electroanal. Chem.* **2018**, *823*, 146–154.
147. Kirankumar, R.; Huang, W.-C.; Chen, H.-F.; Chen, P.-Y. Electropolymerization and characterization of carbazole substituted viologen conducting polymers: The effects of electrolytes and potential applications of the polymers. *J. Electroanal. Chem.* **2018**, *826*, 198–206.
148. Kotanen, C.N.; Karunwi, O.; Alam, F.; Uyehara, C.F.; Guiseppi-Elie, A. Fabrication and in vitro performance of a dual responsive lactate and glucose biosensor. *Electrochim. Acta* **2018**, *267*, 71–79.
149. Phonklam, K.; Wannapob, R.; Sriwimol, W.; Thavarungkul, P.; Phairatana, T. A novel molecularly imprinted polymer PMB/MWCNTs sensor for highly-sensitive cardiac troponin T detection. *Sens. Actuators B Chem.* **2020**, *308*, 127630.
150. Liu, S.; Su, W.; Li, Y.; Zhang, L.; Ding, X. Manufacturing of an electrochemical biosensing platform based on hybrid DNA hydrogel: Taking lung cancer-specific miR-21 as an example. *Biosens. Bioelectron.* **2018**, *103*, 1–5.
151. Maiga, M.; Yazgan, I.; Kariuki, V.M.; Demirkol, D.O.; Sadik, O.A.; Timur, S. Brilliant green sequestered poly (amic) acid film for dual-mode detection: Fluorescence and electrochemical enzymatic biosensor. *Sens. Actuators B Chem.* **2018**, *256*, 71–78.
152. Mattoussi, M.; Matoussi, F.; Raouafi, N. Non-enzymatic amperometric sensor for hydrogen peroxide detection based on a ferrocene-containing cross-linked redox-active polymer. *Sens. Actuators B Chem.* **2018**, *274*, 412–418.
153. Chen, Y.; Liu, X.; Wu, T.; Hou, W.; Liu, M.; Zhang, Y.; Yao, S. Enhanced electrochemical sensitivity towards acetaminophen determination using electroactive self-assembled ferrocene derivative polymer nanospheres with multi-walled carbon nanotubes. *Electrochim. Acta* **2018**, *272*, 212–220.
154. Albayati, S.A.R.; Kashanian, S.; Nazari, M.; Rezaei, S. Novel fabrication of a laccase biosensor to detect phenolic compounds using a carboxylated multiwalled carbon nanotube on the electropolymerized support. *Bull. Mat. Sci.* **2019**, *42*, 187.
155. Hwang, D.W.; Lee, S.; Seo, M.; Chung, T.D. Recent advances in electrochemical non-enzymatic glucose sensors—A review. *Anal. Chim. Acta* **2018**, *1033*, 1–34.

156. Lamiri, L.; Belgherbi, O.; Dehchar, C.; Laidoudi, S.; Tounsi, A.; Nessark, B.; Habelhames, F.; Hamam, A.; Gourari, B. Performance of polybithiophene-palladium particles modified electrode for non-enzymatic glucose detection. *Synth. Met.* **2020**, *266*, 116437.
157. Ayranci, R.; Torlak, Y.; Ak, M. Non-Enzymatic Electrochemical detection of glucose by mixed-valence cobalt containing keggin polyoxometalate/multi-walled carbon nanotube composite. *J. Electroanal. Chem.* **2019**, *166*, B205.
158. Dai, H.; Cao, P.; Chen, D.; Li, Y.; Wang, N.; Ma, H.; Lin, M. Ni-Co-S/PPy core-shell nanohybrid on nickel foam as a non-enzymatic electrochemical glucose sensor. *Synth. Met.* **2018**, *235*, 97–102.
159. Kim, D.-M.; Moon, J.-M.; Lee, W.-C.; Yoon, J.-H.; Choi, C.S.; Shim, Y.-B. A potentiometric non-enzymatic glucose sensor using a molecularly imprinted layer bonded on a conducting polymer. *Biosens. Bioelectron.* **2017**, *91*, 276–283.
160. Wang, J.; Wang, M.; Guan, J.; Wang, C.; Wang, G. Construction of a non-enzymatic sensor based on the poly (o-phenylenediamine)/Ag-NPs composites for detecting glucose in blood. *Mat. Sci. Eng. C* **2017**, *71*, 844–851.
161. Hocevar, M.A.; Fabregat, G.; Armelin, E.; Ferreira, C.A.; Alemán, C. Nanometric polythiophene films with electrocatalytic activity for non-enzymatic detection of glucose. *Eur. Polym. J.* **2016**, *79*, 132–139.
162. Ghanbari, K.; Babaei, Z. Fabrication and characterization of non-enzymatic glucose sensor based on ternary NiO/CuO/polyaniline nanocomposite. *Anal. Biochem.* **2016**, *498*, 37–46.
163. Shahnavaz, Z.; Lorestani, F.; Alias, Y.; Woi, P.M. Polypyrrole–$ZnFe_2O_4$ magnetic nano-composite with core–shell structure for glucose sensing. *Appl. Surf. Sci.* **2014**, *317*, 622–629.
164. Meng, F.; Shi, W.; Sun, Y.; Zhu, X.; Wu, G.; Ruan, C.; Liu, X.; Ge, D. Nonenzymatic biosensor based on CuxO nanoparticles deposited on polypyrrole nanowires for improving detection range. *Biosens. Bioelectron.* **2013**, *42*, 141–147.
165. Çiftçi, H.; Tamer, U. Functional gold nanorod particles on conducting polymer poly (3-octylthiophene) as non-enzymatic glucose sensor. *React. Funct. Polym.* **2012**, *72*, 127–132.
166. Duan, X.; Liu, K.; Xu, Y.; Yuan, M.; Gao, T.; Wang, J. Nonenzymatic electrochemical glucose biosensor constructed by NiCo2O4@ Ppy nanowires on nickel foam substrate. *Sens Actuators B Chem.* **2019**, *292*, 121–128.
167. Kailasa, S.; Geeta, B.; Jayarambabu, N.; Reddy, R.K.K.; Sharma, S.; Rao, K.V. Conductive polyaniline panosheets (CPANINS) for a non-enzymatic glucose sensor. *Mat. Lett.* **2019**, *245*, 118–121.
168. Mohajeri, S.; Dolati, A.; Yazdanbakhsh, K. Synthesis and characterization of a novel non-enzymatic glucose biosensor based on polyaniline/zinc oxide/multi-walled carbon nanotube ternary nanocomposite. *J. Electrochem. Sci. Eng.* **2019**, *9*, 207–222.
169. Kailasa, S.; Rani, B.G.; Reddy, M.S.B.; Jayarambabu, N.; Munindra, P.; Sharma, S.; Rao, K.V. NiO nanoparticles-decorated conductive polyaniline nanosheets for amperometric glucose biosensor. *Mat. Chem. Phys.* **2020**, *242*, 122524.
170. The Nobel Prize in Chemistry 2000. Available online: https://www.nobelprize.org/prizes/chemistry/2000/summary/ (accessed on 27 August 2020).
171. Jia, X.; Ge, Y.; Shao, L.; Wang, C.; Wallace, G.G. Tunable conducting polymers: Toward sustainable and versatile batteries. *ACS Sustain. Chem. Eng.* **2019**, *7*, 14321–14340.
172. Guo, X.; Facchetti, A. The journey of conducting polymers from discovery to application. *Nat. Mater.* **2020**, *19*, 922–928.
173. Heyrovsky, J. The Nobel Prize in Chemistry 1959 Was Awarded to Jaroslav Heyrovsky for His Discovery and Development of the Polarographic Methods of Analysis. Available online: https://www.nobelprize.org/prizes/chemistry/1959/summary/ (accessed on 28 August 2020).
174. Marcus, R.A. The Nobel Prize in Chemistry 1992 Was Awarded to Rudolph A. Marcus for His Contributions to the Theory of Electron Transfer Reactions in Chemical Systems. Available online: https://www.nobelprize.org/prizes/chemistry/1992/summary/ (accessed on 28 August 2020).
175. Moorthy, M.; Kumar, V.B.; Porat, Z.E.; Gedanken, A. Novel polymerization of aniline and pyrrole by carbon dots. *New J. Chem.* **2018**, *42*, 535–540.
176. Maruthapandi, M.; Saravanan, A.; Luong, J.H.T.; Gedanken, A. Antimicrobial properties of polyaniline and polypyrrole decorated with zinc-doped copper oxide microparticles. *Polymers* **2020**, *12*, 1286.
177. Maruthapandi, M.; Saravanan, A.; Luong, J.H.T.; Gedanken, A. Antimicrobial properties of the polyaniline composites against *Pseudomonas aeruginosa* and *Klebsiella pneumoniae*. *J Funct. Biomat.* **2020**, *11*, 59.

178. Cho, N.H.; Shaw, J.E.; Karuranga, S.; Huang, Y.; da Rocha Fernandes, J.D.; Ohlrogge, A.W.; Malanda, B.; Atlas, I.D. Global estimates of diabetes prevalence for 2017 and projections for 2045. *Diabetes Res. Clin. Pract.* **2018**, *138*, 271–281.
179. Villena Gonzales, W.; Mobashsher, A.T.; Abbosh, A. The progress of glucose monitoring—A review of invasive to minimally and non-invasive techniques, devices and sensors. *Sensors* **2019**, *19*, 800.
180. Jo, C.; Ahn, H.; Son, J.; Lee, J.; Byun, M. Packaging and irradiation effect on lipid oxidation, color, residual nitrite content, and nitrosamine formation in cooked pork sausage. *Food Control* **2003**, *14*, 7–12.
181. Uzer, A.; Saglam, S.; Can, Z.; Ercag, E.; Apak, R. Electrochemical determination of food preservative nitrite with gold nanoparticles/p-aminothiophenol-modified gold electrode. *Int. J. Mol. Sci.* **2016**, *17*, 1253.
182. Wrobel, A.G.; Benton, D.J.; Xu, P.; Roustan, C.; Martin, S.R.; Rosenthal, P.B.; Skehel, J.J.; Gamblin, S.J. SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects. *Nat. Struct. Mol. Biol.* **2020**, *27*, 763–767.
183. Lee, J.; Kim, J.; Hyeon, T. Recent Progress in the synthesis of porous carbon materials. *Adv. Mater*. **2006**, *18*, 2073–2094.
184. Song, Y.; Su, D.; Shen, Y.; Gong, C.; Song, Y.; Wang, L. Nitrogen-doped carbon foam as an efficient enzymatic biosensing platform for glucose sensing. *Anal. Methods* **2016**, *8*, 4547–4553.
185. Chen, J.; Zhu, R.; Huang, J.; Zhang, M.; Liu, H.; Sun, M.; Wang, L.; Song, Y. A glucose biosensor based on glucose oxidase immobilized on three-dimensional porous carbon electrodes. *Analyst* **2015**, *140*, 5578–5584.
186. Thakur, B.; Guo, X.; Chang, J.; Kron, M.; Chen, J. Porous carbon and Prussian blue composite: A highly sensitive electrochemical platform for glucose biosensing. *Sens. Bio-Sens. Res*. **2017**, *14*, 47–53.
187. He, X.; Zhou, L.; Nesterenko, E.P.; Nesterenko, P.N.; Paull, B.; Omamogho, J.O.; Glennon, J.D.; Luong, J.H.T. Porous graphitized carbon monolith as an electrode material for probing direct bioelectrochemistry and selective detection of hydrogen peroxide. *Anal. Chem*. **2008**, *4*, 2351–2357.
188. Cui, H.-F.; Du, L.; Guo, P.-B.; Zhu, B.; Luong, J.H.T. Controlled modification of carbon nanotubes and polyaniline on macroporous graphite felt for high-performance microbial fuel cell anode. *J. Power Sources* **2015**, *283*, 46–53.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 5, Issue 3, ISSN No. 2455-2143, Pages 661-670
Published Online July 2020 in IJEAST (http://www.ijeast.com)



RISK AND PREVENTING OF COVID-19 IN A COMMERCIAL AIRCRAFT CABIN: AN OVERVIEW

Chanfiou Ahmed Mboreha
Aeronautical Engineering Department,
Nanjing University of Aeronautics and Astronautics,
Jiangsu, Nanjing, China

Chaharzad Said Abdallah
Second Hospital of Jilin University,
Jilin, Changchun, China

Gaurav Kumar
Department of Mechanical Engineering,
Delhi Technological University, Delhi, India

Abstract **- Aircraft passengers and crew are lining a coronavirus disease pandemic. This pandemic may last for many months, stressing the aviation industries. With over two billion airline passengers yearly, in-flight transmission of contagious diseases is an essential worldwide health interest. Some attitudes of in-flight transmission have been registered, cases of coronavirus disease (COVID-19) have been reported in all over the world, traveling raises your risks of receiving and scattering COVID-19. Because of how air flows and is cleaned on airplanes, maximum viruses and other microorganisms do not spread effortlessly on flights. Nevertheless, there may be a possibility of receiving COVID-19 on crowded flights if there are other passengers inside the aircraft cabin with COVID-19. So the main objective of this study is to explain the relation between the aircraft cabin ventilation systems with the spread of COVID-19 inside an aircraft cabin. Aircraft ventilation strategies and their effect on air quality were discussed in this study. Control measures and some aircraft cabin seats arrangement proposals were also presented in this study.**

Keywords- **COVID-19**, **aircraft ventilation system**, **airborne disease**, **aircraft cabin**.

I. INTRODUCTION

The impact of air travel on the expansion of infectious diseases has led to significant fear but limited study (Mangili et al 2005; Vogt TM ,2006). There is a stat which says that about 1 billion people travel in airplanes per year. In reality, more travelers are likely to be scared of any type of microbe diseases which is infectious in nature on their journey. There are many important ways in which air travel can be affected the worldwide expansion of new pop up and pre-existing infectious disease. Transmissibility may be expanded on the aircraft via close contact; airborne expand itself through small-particle aerosols, as in the case of some acute respiratory syndrome (SARS). Our understanding of the spread of inhalation infections on aircraft is very imperfect. Key questions include how often dispersal can occur by contaminated surfaces, large droplet spread, and airborne spread of small-particle aerosols, as well as the risk for specific organisms. The scientific uncertainty limits the capability to produce any preventive measures. Many types of commercial aircraft use vertical airflow and high-efficiency particulate air (HEPA) filters that should limit the exposure to small airborne droplets. However, there are no regulations requiring HEPA filters or for testing of the function of filters. (Ozonoff D, 2005)

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 5, Issue 3, ISSN No. 2455-2143, Pages 661-670
Published Online July 2020 in IJEAST (http://www.ijeast.com)



The COVID-19 pandemic presents specific dangers that encounter airline travelers. Travelling contains more than a few individual connections and sitting close to others for long times, both in flight and at airport entrances. These activities can raise the accidental contact to transmissible illnesses. This review explained the relation between the aircraft cabin ventilation systems with the extent of COVID-19 inside an aircraft cabin. Aircraft ventilation approaches and their effect on air quality were discussed in this study. Control measures and some aircraft cabin seats arrangement proposals are also presented in this study.

II. PURPOSE

The main Purpose of this study is to explain the relation between the aircraft cabin ventilation systems with the spread of COVID-19 inside an aircraft cabin. Aircraft ventilation strategies and their effect on air quality were discussed in this study. Control measures and some aircraft cabin seats arrangement proposals were also presented in this study.\

III. LITERATURE REVIEW

On 12 January 2020, the World Health Organization (WHO) confirmed that a novel coronavirus was the cause of a respiratory illness in a cluster of people in Wuhan City, Hubei Province, China, which was reported to the WHO on 31 December 2019 (Reynolds Matt,2020). The case fatality ratio for COVID-19 has been much lower than SARS of 2003 (GOV.UK report,2020) but the transmission has been considerably greater, with a significant total death toll (WASAHQ Report , 2020) . COVID-19 is mainly initiated by contact or droplet transmission attributed to relatively large respiratory particles which are matter to gravitational forces and travel only approximately one meter from the patient. Airborne transmission may occur if patient respiratory activity or medical procedures generate respiratory aerosols. These aerosols contain particles that may float much longer distances and remain airborne longer, but their infective power is uncertain. Aircraft passengers and crew are confronting a coronavirus disease pandemic. This pandemic may last for many months, stressing the aviation industries.

Keeping Aircraft passengers and crew safe, healthy, and available to travel throughout this pandemic is critical. The consistent use of appropriate personal protective equipment (PPE) will help the assurance of its availability and Aircraft passenger's protection (Livingston E, Desai A & Berkwits M.,2020) Aircraft passengers and crew are at risk of contagion during the severe acute respiratory syndrome coronavirus-2 pandemic. International guidance suggests direct droplet transmission is likely and airborne transmission occurs only with aerosol-generating procedures (Wilson NM, Norton A, Young FP & Collins DW, 2020)**.** Aircraft passengers and crew should appraise the current evidence regarding spread and apply this to the contagion occurrence. Measures to soften airborne transmission should be employed at times of risk. However, the mechanisms and risk factors for transmission are largely unconfirmed. Whilst awaiting robust evidence, a precautionary approach should be considered to assure Aircraft passengers and crew safety.

1- Aircraft cabin environment

Commercial airplanes voyage at a typical altitude of 11,000 meter where the outside temperature is about -55 ℃, the atmospheric pressure is only about one-fifth of that at the sea level and the relative humidity is near zero. Under such a risky ambient environment, human beings cannot survive without being protected by an environmental control system (ECS) (Elmaghraby, H.A., Y.W. Chiang, and A.A. Aliabadi , 2018 ; Barnes, N.M., et al 2018) . As a part of the ECS, air distribution system plays roles including distributing air suitably, cleaning up filthy air, and diminishing cross disease infection, towards to creating relaxed, healthy and safe cabin surroundings for passengers and crews. However, the current widely-used mixing air distribution system promotes air mixing by supplying high momentum air at the ceiling level and exhausting contaminated air at the deck level. The widespread air mixing is anticipated to dilute cabin inside generated contaminants and thus may impose risks of infectious airborne disease transmission. The outbreaks of airborne communicable diseases, such as the severe acute respiratory syndrome (SARS) in 2003 and COVID-19 in 2019, confirm such risks are not imaginary but may come to realism (Lidia Morawska, Junji Cao, 2020). In addition, many long criticisms about cabin air quality and thermal ease should also be put with significant attentions. To demonstrate how the ECS works, a volume of air will be followed as it flows continuously from outside the airplane into the airplane cabin and, finally, exits the airplane. Figure 1 is a schematic of the ECS components for the Boeing 767 airplane. The Boeing 767 is a modern generation airplane that uses 50% filtered recirculated air and 50% outside air (a typical of modern generation airplanes).

The air delivered to passengers and crew on commercial jet aircraft is naturally a blend of conditioned external (ambient) air that has been diverted to the cabin from the engine compressor stages, and air that is taken from the cabin, filtered, and recirculated. The first filter of recirculated air traps the largest particles. Then, on most modern aircraft, the air passes through high efficiency particular air (HEPA) filters before reentering the passenger cabin (Mangili A, Vindenes T, Gendreau M., 2015). The most efficient HEPA filters capture 99.97% of particles between 0.1 and 0.3 micrometer and 100% of the larger particles. In a count, the ultra-small contagion also gets captured by Brownian and electrostatic forces. The environmental control system is designed to minimize the introduction of harmful contaminants into the cabin and to control cabin pressure, ventilation, and temperature. Most commercial aircraft provide 10 to 15 air changes per hour of the passenger cabin depending upon the aircraft type, as the original cabin air is diluted with incoming ambient air. Cabin

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 5, Issue 3, ISSN No. 2455-2143, Pages 661-670
Published Online July 2020 in IJEAST (http://www.ijeast.com)



airflow is mainly side to side and compartmentalized into four to seven seat rows within the passenger cabin, thereby limiting longitudinal (front to back and vice versa) cabin contamination. However, although the principle is side-to-side airflow, evidence employing computational fluid dynamic modeling in a mockup full-sized aircraft cabin section suggests that airflow patterns can be influenced by many factors including seat and cabin geometry, occupancy density, and thermal loads of passengers and equipment. Individual risk of infection of COVID-19 within the aircraft cabin, involves many factors, including chance, source strength (generation rate of infectious agent), exposure (proximity to the source and duration of the exposure), and cabin ventilation. Infections during flights can be spread not only by aerosols that remain airborne and can be inhaled, but also by large droplets that settle on surfaces or by direct contact with secretion, body fluids, or contaminated surfaces.

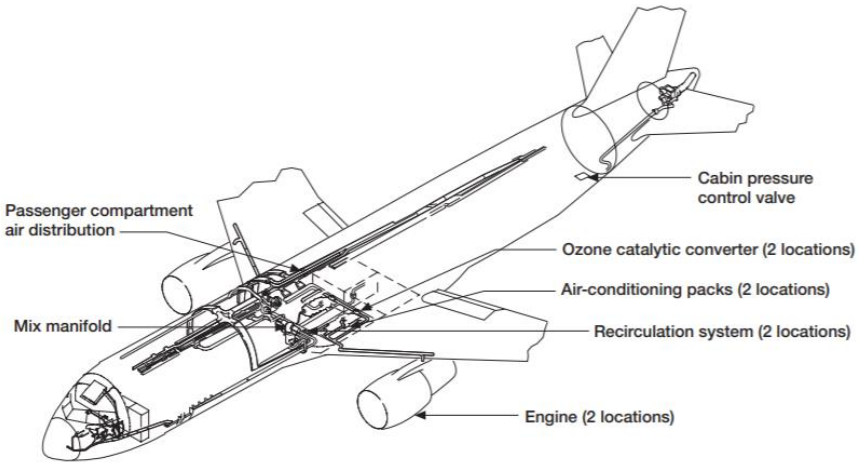


Figure 1: System layout for the Boeing 767 ECS (Hunt, E.H., et al,1995)

2- Aircraft ventilation strategies and their effect on air quality and COVID-19 infection aboard

Different ventilation plans have been examined as measures to control infection transmitted via airborne route and other gaseous and particulate contaminants commonly found in the aircraft cabin envelop. For aircraft cabins, ventilation effectiveness is defined as the portion of the outside fresh air supplied to the cabin that reaches the breathing zone of the occupants. Major studies that examined the effect of one or more ventilation system, such as mixing, displacement, and personalized ventilation, on air quality and disease spread aboard aircraft are discussed in different researches. Under-aisle and underfloor ventilation systems are usually used as other names for the displacement ventilation system aboard aircraft. In such systems, air is supplied vertically from an under-floor plenum, usually located under the aisles, through perforated or nozzle vents. In aircraft cabins, personalized ventilation has been employed in various forms as a measure to protect the micro-environment of occupants from the invasion of contaminants and infectious organisms. (Wilson NM, Norton A, Young FP & Collins DW ,2020)

Some researches confirmed that inlet and outlet should be near to each other to decrease airborne transmission but they did not explain the effect of inlet and outlet location on airflow distribution and thermal stability. To reduce airborne transmission, path transfer by air should be lowest and re-circulation should be less. Credible scientific investigations of cabin air quality have been conducted by the National Academy of Sciences, United States Department of Transportation (DOT), National Institute for Occupational Safety and Health (NIOSH), Centers for Disease Control and Prevention (CDC), independent research groups, and airplane manufacturers (Wilson, N. M. Norton, A.Young, F. P.Collins, D. W,2020). Numerous engineers and researchers have worked on improving cabin air quality, thermal comfort, and air flow in recent years. As stated before, the ECSs function is to supply a given amount of fresh air at a certain pressure and temperature in order to maintain the cabin at such conditions that the comfort of passengers and crew is guaranteed. It is essential that ECSs in an aircraft be proficient of quickly heating up or cooling down cabins in order to deliver an acceptable thermal environment for passengers and crew members (Carlo Giaconia, Aldo Orioli, Alessandra Di Gangi , 2013 examined the temperature, pressure and humidity of an indoor air in aircraft.)

Some other studies have been published comparing displacement ventilation (DV) with other ventilation systems. The DV is an air distribution system which delivers cool fresh air with low velocity, generally at floor level and eliminates exhausted air at the ceiling. The temperature distribution is

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 5, Issue 3, ISSN No. 2455-2143, Pages 661-670
Published Online July 2020 in IJEAST (http://www.ijeast.com)



described by stratification: while cool, fresh and clean air can be found at floor level, the contaminated air rises up to the ceiling and is extracted there efficiently improving the air quality. **(**Zhu S., Srebic J ,2012 analyzed the performance of DV and Mixing Ventilation (MV) in different kinds of rooms). Computational Fluid Dynamics (CFD) models were used to compare DV and MV regarding airflow, temperature and comfort indices. DV turned out to provide a comparable or even better comfort level than MV, except for the space in the vicinity of floor diffusers or major heat sources. (T Zhang*, Q Chen, 2005 Analyzed DV, MV and personalized ventilation in a section of a Boeing 767 cabin using CFD modelling). Temperature stratification was found in both systems but the vertical gradient was less than 3 K. While in MV, high air velocities (and CO2 concentrations) were found, DV provided more comfort.

Several research work anxious compared DV, MV and PV(Personal Ventilation) with lateral provision in a mock-up based on different Airbus designs (Tengfei Zhang, Qingyan (Yan) Chen ,2007) . Air flow and temperature measurements were made as well as experiments with test persons and CFD simulations. The vertical temperature difference in DV and in MV with lateral support was in line with the standards for a volumetric flow rate of 77.8 l/s/cabin and low air velocities were measured (< 0.2 m/s) (Julia Maier, Claudia Marggraf-Micheel, Tobias Dehne, Johannes Bosbach, 2017). An investigation conducted by the Committee on Airliner Cabin Air Quality in 1986 showed that the relative humidity (RH) decreases from 23% to 2% under regular cabin ventilation rates (Long Chen, Xingjuan Zhang, Chao Wang, Chunxin Yang ,2017). Low RH in the aircraft cabin may cause a number of problems, such as skin dryness, eye irritation, increased static electricity, and increases the survival time of virus such as SARS-CoV-2, influenza and SARS-CoV-1 viruses.

During and after the COVID-19 pandemic, researchers should study a new ventilation system with personalized air supply suitable to build a mini-environment in the commercial aircraft cabin. A personalized air distribution system provides clean air straight to the respiration area of somebody. The system generates a desirable micro-environment with fresh air. The personalized air distribution system may deliver a decent air quality, but the danger could be even higher than the displacement ventilation system. Researchers should study a suitable seat arrangement, which may decrease the risk of infectious diseases like COVID-19. Hence in this study we proposed some seat arrangement concepts which will need experimental and CFD simulation analysis before concluding the suitable seat arrangement in a commercial aircraft cabin.

3- Control measures

A key constituent of infection control, particularly for diseases with airborne spread has been reported. For example, the COVID-19 outburst involved passengers to seat as far as two rows from the index passenger. Multiple mathematical models of probability of infection are available, and perhaps the most commonly used is the Wells Riley calculation, which has been used to study both tuberculosis and measles outbreaks (Noakes CJ, Sleigh PA.,2009) Correct ventilation within any confined space decreases the concentration of airborne organisms in a logarithmic fashion. In the case of recirculated systems, such as on recent aircraft, where several sections of air from the cabin are recirculated and mixed with fresh air, this relationship holds only if the recirculated air experiences filtration through HEPA filters. Although the practice of removing transmissible airborne particles from recirculated air through HEPA filtering is well accepted as an impressive transmission control measure, the use of HEPA filters within commercial aircraft is neither required nor regulated by the Federal Aviation Administration or its British (Civil Aviation Authority) and European (Joint Aviation Authority) counterparts (Mangili A, Vindenes T, Gendreau M.,2015)

A 2004 survey of major U.S. commercial air carriers by the U.S. Government Accountability Office found that 15% of large commercial aircraft within the U.S. domestic fleet did not routinely use HEPA air filtering and that nearly 50% of smaller regional commercial aircraft did not (David Schlossberg, 2016). Incorporating epidemiological data into risk assessment mathematical models may provide insight into how proximity and ventilation influence disease transmission aboard commercial aircraft. A major public health situation is created by the expectation of travel from one country to another within one day , a period of time which is little than the incubation time period. This allows for most passengers infected with COVID-19 to pass through entry screening undetected and reach their destination before they become symptomatic. Although many countries practice entry screening, studies reveal that this is very ineffective. Khan et al.'s retrospective evaluation of entry and exit screening of airline travelers during the 2009–2010 H1N1 pandemic states that entry screening is an anachronism and says interventions to mitigate the international spread of infectious disease whether through alertness or response would have the greatest global impact if implemented as close as possible to the sources of any future epidemic threats (i.e., exit screening).(Khan K, Eckhardt R, Brownstein JS, Naqvi R, Hu W, Kossowsky D, Scales D, Arino J, MacDonald M, Wang J, Sears J &Cetron MS, 2013)

4- Two rows arrangement

Any passenger flight from any country infected by the COVID-19, who was seated within two rows of a passenger diagnosed with COVID-19 shortly after they had taken a flight from one region to another should also be quarantined for fourteen days. But the statement led some to wonder why only those seated within two rows of the passenger; the answer is "That is close contact within the context of a plane." This approach is based on guidance from the World Health Organization (WHO), which defines "close contact" as being within one meter of an infectious passenger. Outside the situation of a plane, close

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 5, Issue 3, ISSN No. 2455-2143, Pages 661-670
Published Online July 2020 in IJEAST (http://www.ijeast.com)



contact is defined as "face-to-face contact" or "spending more than 15 minutes within two meters of an infected person.
This review devised a "seating heat map", to show the chance of every traveler on a flight becoming infected by somebody with COVID-19 who is seated in an aisle seat in the middle of the plane. The imaginary infectious passenger is seated at the red seat. Passengers seated within a one meter radius of the putative patient were at highest risk (yellow seats), followed by passenger seated within more than one meter radius(Blue seats) were at low risk , followed by a crew passing in front of the putative patient was at high risk , another crew member was passing within more than one meter radius in front the sick passenger was at medium risk and finally a crew member was passing within more than one meter radius behind the sick passenger was a low risk. Public health bodies making contact with passenger's seated two rows in either direction should, in normal circumstances, be more than careful enough. This review mapped the behavior of passengers seated in economy class on a Boeing 767 cabin. Our study advises that the aisle seats is the riskiest place to sit, even someone in an aisle seat just two rows in front of an infected passenger would have bit to worry about. It is not known yet precisely how this novel of coronavirus (COVID-19) is spread.

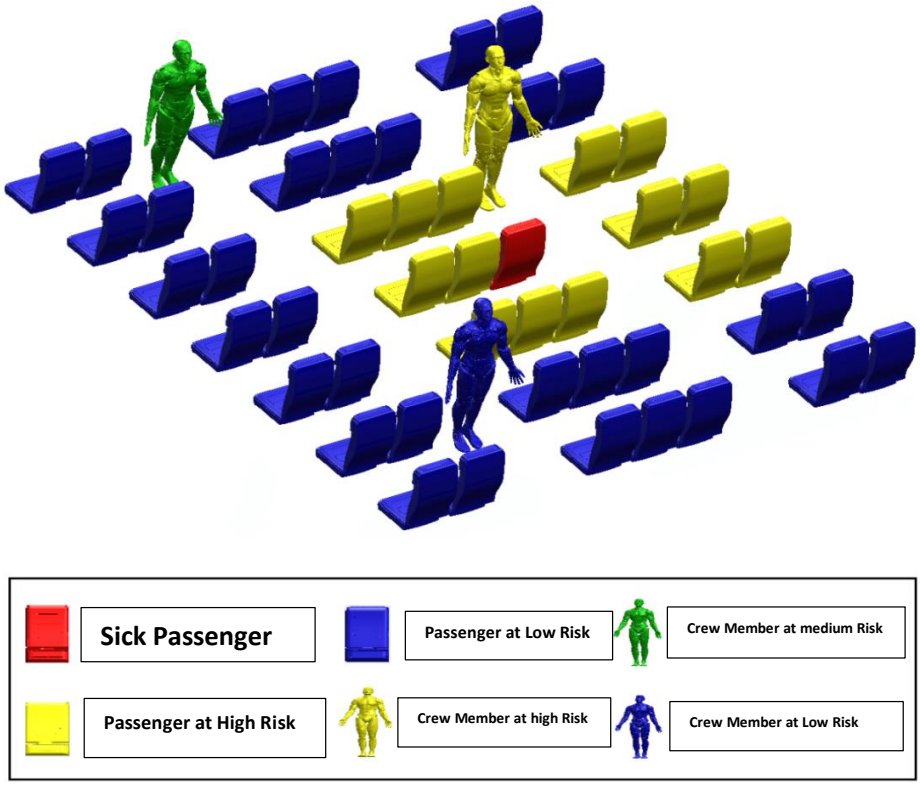


Figure 2: The 2-row infection zone diagram

5- Seat arrangement Proposals for preventing the Spread of COVID-19 in aircraft cabins

The passenger seat arrangement for the aircraft has been the same since years of aviation flights. Due to the beginning of various diseases and their spread in the aircraft cabin, it is of interest for the researcher to try to investigate different seat arrangement, which may contribute in decrease level of infection for the passengers (due to the sick passenger). Therefore, the following three arrangements of seats are proposed, that may be numerically simulated for the investigation of infection risk.

Firstly, we designed a concept that features forward-facing window and aisle seats with rear-facing middle seats as can be seen in figure 3. The middle seats are flanked by a wraparound clear shield that separates them from their two neighboring seats. This is just one of many designs we have seen in the wake of COVID-19 as the travel industry prepares for a more cautious approach as things start up again. This concept is renowned by the opposite position of the middle chair of the triple to guarantee the extreme separation between travelers

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 5, Issue 3, ISSN No. 2455-2143, Pages 661-670
Published Online July 2020 in IJEAST (http://www.ijeast.com)

seated next to each other. Even though traveler seated on the side chairs, aisle and fuselage, stay to be positioned in the flight direction as normal, the passenger sitting in the middle is facing backwards. In fact this arrangement permits all three travelers to be parted with a protection made of translucent material that separates them from each other, making a shielding barrier from each person; each traveler has its personal space quarantined from others.

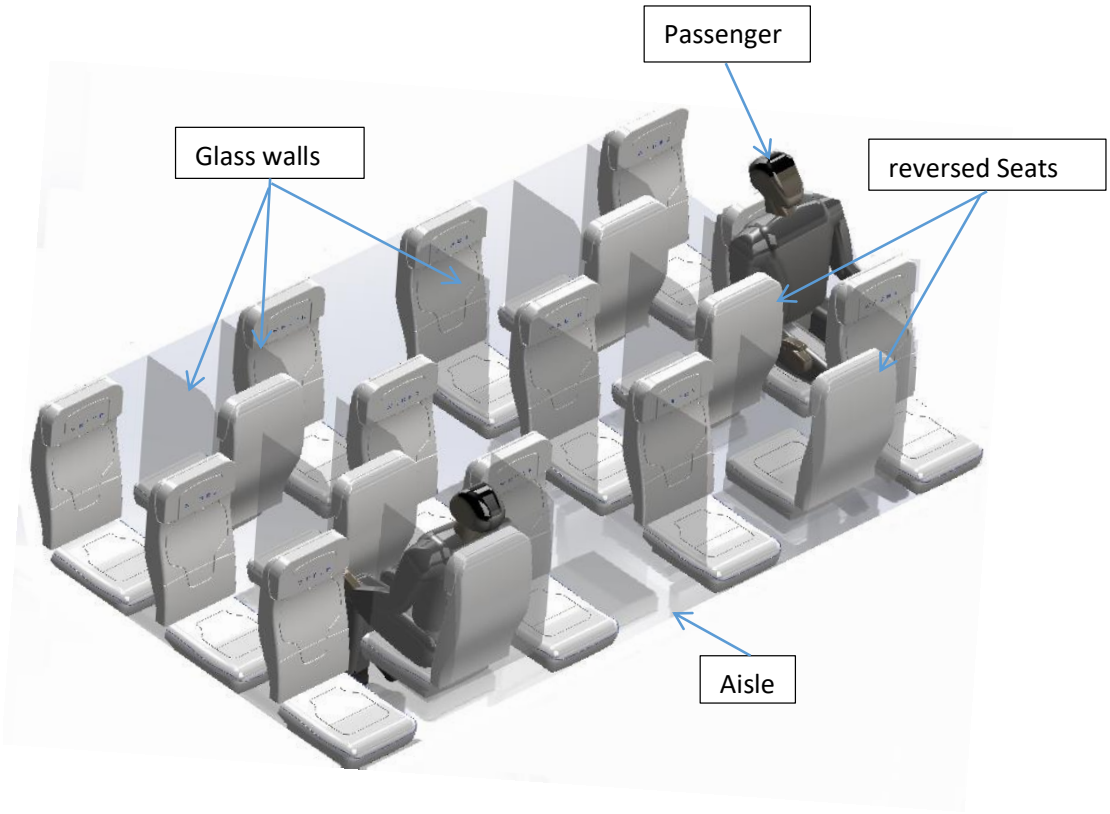


Figure 3 : Forward-facing window and aisle seats with rear-facing middle seats Concept

Secondly, we designed a seat arrangement concepts, which we use the Boeing 737 cabin seat arrangement modification as can be seen in figure 4, which the second seat arrangement concept , it is assumed that after the take-off the seats are rotated 90 degrees (towards the window) .

Each of the rows is again divided with a partition wall. There are personal vents for each passenger (from the top), along with the Longitudinal Inlet present in the top. There are also longitudinal Outlets present on each side of the cabin.

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 5, Issue 3, ISSN No. 2455-2143, Pages 661-670
Published Online July 2020 in IJEAST (http://www.ijeast.com)

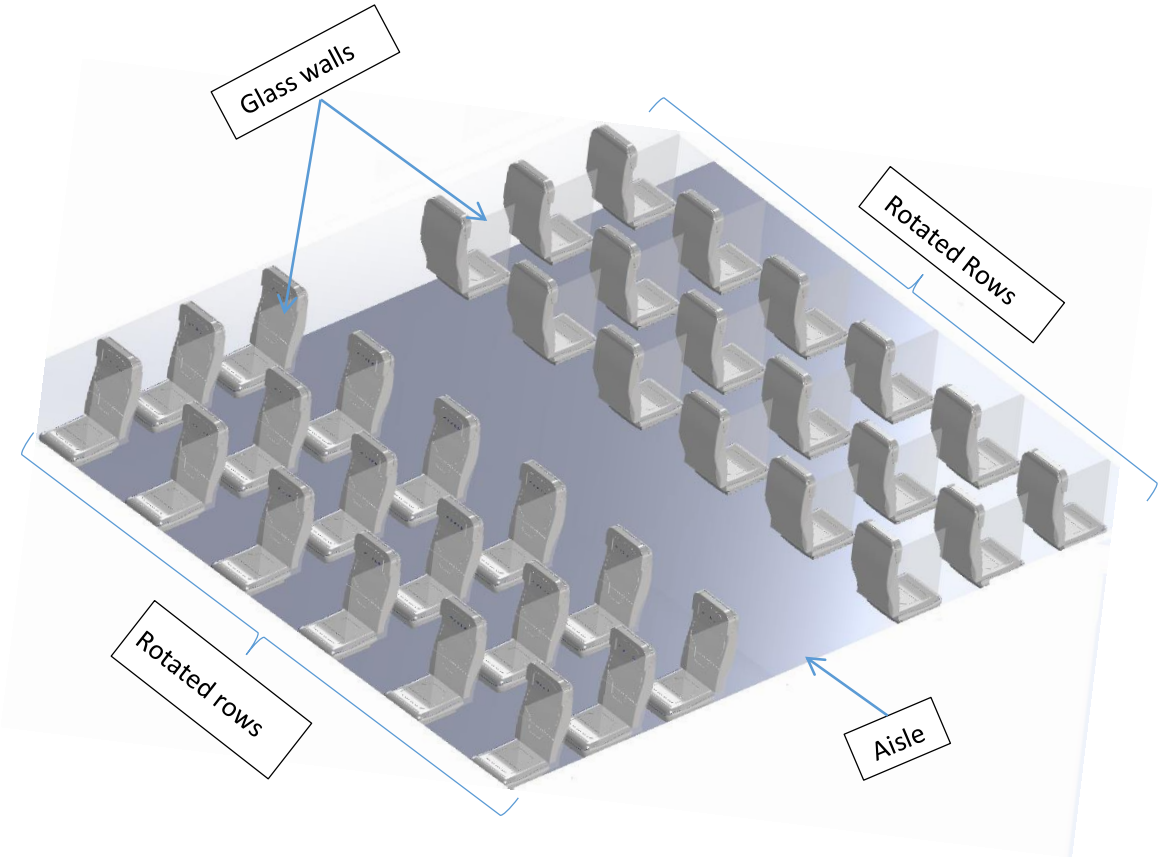


Figure 4 : The rotation of rows arrangement concept

Finally, we designed a third seat arrangement concept, it is assumed that after the take-off the seats are rotated 90 degrees in the pattern seen in figure 5. The middle row is with a partition wall, for preventing them to be in the direct vicinity of the second column passenger. There are personal vents for each passenger (from the top), along with the Longitudinal Inlet present in the top. There are two longitudinal Outlets present on each side of the cabin along with the horizontal outlets present after each row.

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 5, Issue 3, ISSN No. 2455-2143, Pages 661-670
Published Online July 2020 in IJEAST (http://www.ijeast.com)

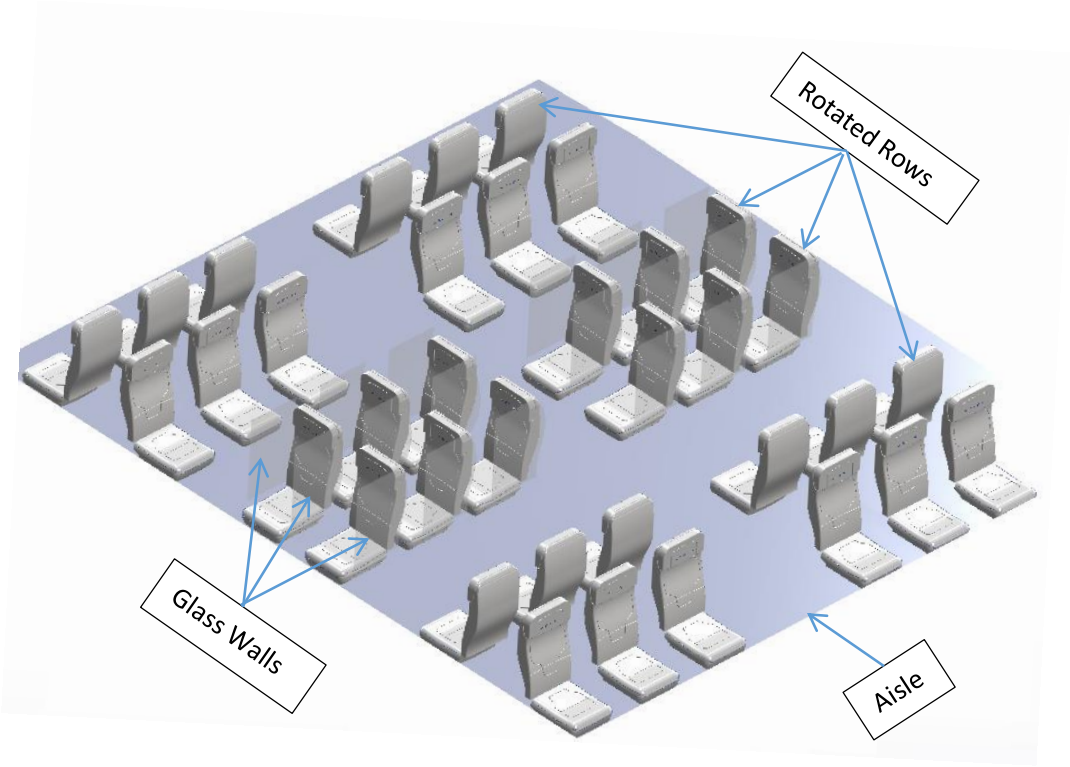


Figure 5 : Multiple rotations of seats arrangement concept

IV. METHODOLOGY

The authors are all contributed for this review. After doing some researches about the risk infection of COVID-19 in an enclosed room or enclosed space, the authors decided to explain the infection risk of COVID-19 inside an aircraft cabin. We came out with some control measures for the decrease of the infection risk of COVID-19 inside an aircraft cabin. As we know that the passenger seat arrangement for the aircraft has been the same since years of aviation flights and due to the beginning of COVID-19 and its spread in the aircraft cabin, it is of interest for the researcher to try to investigate different seat arrangement, which may contribute in decrease level of infection for the passengers (due to the sick passenger). Therefore, the authors proposed three arrangements of seats that may be numerically simulated for the investigation of infection risk. The authors used Solidworks Software to Build the seat arrangements.

V. CONCLUSIONS

The way to destroy almost any risk of infection in an aircraft cabin and to prevent the aircraft from serving as a transport for international epidemic spreading is to stop travelers from traveling if they either have been exposed considerably to a contagious disease like COVID-19. Prevention is the most essential way of control and requires an active movement. For aircraft passengers, practicing hand hygiene and wearing surgical masks remains the most efficient way of decreasing probability of transmission. Developed international regulations respecting the examination, authorization, and reparation of aircraft environmental control systems are secured. Regulations necessitate HEPA filters for any aircraft using recirculated air are needed to decrease the probability of infectious spread on and by commercial aircraft. Due to the beginning of COVID-19 and its spread in the aircraft cabin, it is of interest for the researcher to try to investigate different seat arrangement, which may contribute in decrease level of infection for the passengers. Our research advises that the aisle seats is the riskiest place to sit, even someone in an aisle seat



just two rows in front of an infected passenger would have bitty to worry about. Aviation industries should study a novel personal ventilation suitable to build a mini-environment inside a commercial aircraft cabin.

VI. RECOMMENDATIONS

1. Aviation industries should study a novel personal ventilation suitable to build a mini-environment inside a commercial aircraft cabin
2. A novel seat arrangement may also be studied for future aircraft cabins.

VII. ACKNOWLEDGEMENTS

Authors would like to thank their Professors and colleagues for their guides and supports. Especially Mr. Raj Kumar Singh of Delhi Technological University to support through out the research work & suggesting some good techniques.

VIII. REFERENCES

[1] Elsevier. "Novel Coronavirus Information Center". Elsevier Connect. Archived from the original on 30 January 2020. Retrieved 15 March 2020.

[2] Reynolds, Matt (4 March 2020). "What is coronavirus and how close is it to becoming a pandemic?". Wired UK. ISSN 1357-0978. Archived from the original on 5 March 2020. Retrieved 5 March 2020.

[3] Crunching the numbers for coronavirus". Imperial News. Archived from the original on 19 March 2020. Retrieved 15 March 2020

[4] "High consequence infectious diseases (HCID); Guidance and information about high consequence infectious diseases and their management in England". GOV.UK. Archivedfrom the original on 3 March 2020. Retrieved 17 March 2020.

[5] "World Federation Of Societies of Anaesthesiologists – Coronavirus". www.wfsahq.org. Archived from the original on 12 March 2020. Retrieved 15 March 2020.

[6] Livingston E, Desai A, Berkwits M. Sourcing Personal Protective Equipment During the COVID-19 Pandemic. JAMA. 2020 Mar 28. doi: 10.1001/jama.2020.5317. Epub ahead of print. PMID: 32221579.

[7] Wilson NM, Norton A, Young FP, Collins DW. Airborne transmission of severe acute respiratory syndrome coronavirus-2 to healthcare workers: a narrative review. Anaesthesia. 2020 Apr 20:10.1111/anae.15093. doi: 10.1111/anae.15093.

[8] Elmaghraby, H.A., Y.W. Chiang, and A.A. Aliabadi, Ventilation strategies and air quality management in passenger aircraft cabins: A review of experimental approaches and numerical simulations. Science and Technology for the Built Environment, 2018. 24(2): p. 160-175.

[9] Barnes, N.M., et al., In-Cabin Air Quality during Driving and Engine Idling in Air-Conditioned Private Vehicles in Hong Kong. International Journal of Environmental Research and Public Health, 2018. 15(4).

[10] Lidia Morawska, Junji Cao, Airborne transmission of SARS-CoV-2: The world should face the reality, Environment International, Volume 139, 2020, 105730, ISSN 0160-4120, doi.org/10.1016/j.envint.2020.105730.

[11] Mangili A, Vindenes T, Gendreau M. Infectious Risks of Air Travel. Microbiol Spectr. 2015 Oct;3(5). doi: 10.1128/microbiolspec.IOL5-0009-2015. PMID: 26542037.

[12] Hunt, E.H., et al., Commercial airliner environmental control system - Engineering aspects of cabin air quality. Aerospace Medical Association Annual Meeting, 1995: p. 1-8.

[13] Wilson NM, Norton A, Young FP, Collins DW. Airborne transmission of severe acute respiratory syndrome coronavirus-2 to healthcare workers: a narrative review. Anaesthesia. 2020 Apr 20:10.1111/anae.15093. doi: 10.1111/anae.15093. Epub ahead of print. PMID: 32311771; PMCID: PMC7264768.

[14] Wilson, N. M. Norton, A.Young, F. P.Collins, D. W.”Airborne transmission of severe acute respiratory syndrome coronavirus-2 to healthcare workers: a narrative review” Anaesthesia-2020 ;doi.10.1111/anae.15093,PMC7264768

[15] Carlo Giaconia, Aldo Orioli, Alessandra Di Gangi,Air quality and relative humidity in commercial aircrafts: An experimental investigation on short-haul domestic flights, Building and Environment, Volume 67,2013,Pages 69-81,ISSN 0360-1323,doi.org/10.1016/j.buildenv.2013.05.006.

[16] Zhu S., Srebic J., et all, An advanced numerical model for the assessment of airborne transmission of influenza in bus microenvironments, Building and Environment 47, pp. 67-75, 2012

[17] T Zhang*, Q Chen. Comparison of different ventilation systems for commercial. Aircraft cabins. Indoor Air 2005. 3205.

[18] Tengfei Zhang, Qingyan (Yan) Chen,Novel air distribution systems for commercial aircraft cabins, Building and Environment,Volume 42, Issue 4,2007,Pages 1675-1684,ISSN 0360-1323, doi.org/10.1016/j.buildenv.2006.02.014.



[19] Julia Maier, Claudia Marggraf-Micheel, Tobias Dehne, Johannes Bosbach,Thermal comfort of different displacement ventilation systems in an aircraft passenger cabin,Building and Environment,Volume111,2017,Pages256-264,ISSN0360-1323, doi.org/10.1016/j.buildenv.2016.11.017.

[20] Long Chen, Xingjuan Zhang, Chao Wang, Chunxin Yang, A novel environmental control system facilitating humidification for commercial aircraft,Building and Environment,Volume126,2017,Pages34-41,ISSN0360-1323, doi.org/10.1016/j.buildenv.2017.09.013.

[21] Noakes CJ, Sleigh PA. Mathematical models for assessing the role of airflow on the risk of airborne infection in hospital wards. J R Soc Interface. 2009 Dec 6;6 Suppl 6(Suppl 6):S791-800. doi:10.1098/rsif.2009.0305.focus.

[22] Mangili A, Vindenes T, Gendreau M. Infectious Risks of Air Travel. Microbiol Spectr. 2015 Oct;3(5). doi:10.1128/microbiolspec.IOL5-0009-2015.

[23] David Schlossberg ,"Infections of Leisure" Fifth Edition-2016 , eISBN : 9781555819231 , doi:10.1128/9781555819231

[24] Khan K, Eckhardt R, Brownstein JS, Naqvi R, Hu W, Kossowsky D, Scales D, Arino J, MacDonald M, Wang J, Sears J, Cetron MS. Entry and exit screening of airline travellers during the A(H1N1) 2009 pandemic: a retrospective evaluation. Bull World Health Organ. 2013 May 1;91(5):368-76. doi: 10.2471/BLT.12.114777.

CEEW
THE COUNCIL

Centre for
Energy Finance

Scaling Up Solar Manufacturing to Enhance India's Energy Security

Report | August 2020

Rishabh Jain, Arjun Dutt, and Kanika Chawla





Domestic solar module manufacturers import most of the components required for manufacturing, including solar cells.

Image: Vikram Solar ltd

CEEW
THE COUNCIL
Centre for
Energy Finance

Scaling Up Solar Manufacturing to Enhance India's Energy Security

Rishabh Jain, Arjun Dutt, and Kanika Chawla

Report
August 2020
ceew.in

Copyright © 2020 Council on Energy, Environment and Water (CEEW).



Open access. Some rights reserved. This work is licenced under the Creative Commons Attribution- Noncommercial 4.0. International (CC BY-NC 4.0) licence. To view the full licence, visit: www. creativecommons.org/licences/ by-nc/4.0/legalcode.

Disclaimer: The views expressed in this work are those of the authors and do not necessarily reflect the views and policies of the Council on Energy, Environment and Water.

Cover image: Vikram Solar Limited.

Suggested citation: Jain, Rishabh, Arjun Dutt and Kanika Chawla, 2020. *Scaling Up Solar Manufacturing to Enhance India's Energy Security*. New Delhi: Council on Energy, Environment and Water.

Peer reviewers: Deepak Gupta, Senior Vice President, ReNew Power; Paramita Sahoo, Head—Policy Advocacy, Tata Power; Sujoy Ghosh, Vice President—India & Middle East, First Solar; and Arunabha Ghosh, CEO, Council on Energy, Environment and Water.

Publication team: Alina Sen (CEEW), Mihir Shah (CEEW), Venkatesh Krishnamoorthy, Aspire Designs, and Friends Digital.

Acknowledgments: We would like to thank Alpesh Dave, DGM, Goldi Solar Private Limited; Bharat Bhut, Director, Goldi Solar Private Limited; Dinesh Shenoy, VP- Strategy, Waaree Energies Limited; Manish Karna, General Manager–Business Development, Adani Green Energy; Manmeet Singh, Assistant Manager, Vikram Solar Limited; Nandkumar Pai, CEO, Renewsys India Private Limited; Rohit Kumar, Head-India, REC Solar, Sundeep Gupta, MD, Jakson Limited Group, and Vineet Mittal, Director, Navitas Green Solutions Private Limited for sharing qualitative and quantitative information on solar manufacturing process which went a long way in enhancing the report's output.

Organisation: The **Council on Energy, Environment and Water** (ceew.in) is one of Asia's leading not-for-profit policy research institutions. The Council uses data, integrated analysis, and strategic outreach to explain and change the use, reuse, and misuse of resources. It prides itself on the independence of its high-quality research, develops partnerships with public and private institutions and engages with the wider public. In 2020, CEEW has once again been featured across nine categories in the *2019 Global Go To Think Tank Index Report*. It has also been consistently ranked among the world's top climate change think tanks. Follow us on Twitter @CEEWIndia for the latest updates.

Council on Energy, Environment and Water
Sanskrit Bhawan, A-10, Qutab Institutional Area
Aruna Asaf Ali Marg, New Delhi - 110067, India

CEEW Centre for Energy Finance

The CEEW Centre for Energy Finance (CEEW-CEF) is an initiative of the Council on Energy, Environment and Water (CEEW), one of Asia's leading think tanks.

CEEW-CEF acts as a non-partisan market observer and driver that monitors, develops, tests, and deploys financial solutions to advance the energy transition. It aims to help deepen markets, increase transparency, and attract capital in clean energy sectors in emerging economies. It achieves this by comprehensively tracking, interpreting, and responding to developments in the energy markets while also bridging gaps between governments, industry, and financiers.

The need for enabling an efficient and timely energy transition is growing in emerging economies. In response, CEEW-CEF focuses on developing fit-for-purpose market-responsive financial products. A robust energy transition requires deep markets, which need continuous monitoring, support, and course correction. By designing financial solutions and providing near-real-time analysis of current and emerging clean energy markets, CEEW-CEF builds confidence and coherence among key actors, reduces information asymmetry, and bridges the financial gap.

Financing the energy transition in emerging economies

The clean energy transition is gaining momentum across the world with cumulative renewable energy installation crossing 1000 GW in 2018. Several emerging markets see renewable energy markets of significant scale. However, these markets are young and prone to challenges that could inhibit or reverse the recent advances. Emerging economies lack well-functioning markets. That makes investment in clean technologies risky and prevents capital from flowing from where it is in surplus to regions where it is most needed. CEEW-CEF addresses the urgent need for increasing the flow and affordability of private capital into clean energy markets in emerging economies.

CEEW-CEF's focus: analysis and solutions

CEEW-CEF has a twin focus on markets and solutions. CEEW-CEF's market analysis covers energy transition–related sectors on both the supply side (solar, wind, energy storage) and demand-side (electric vehicles, distributed renewable energy applications). It creates open-source data sets, salient and timely analysis, and market trend studies.

CEEW-CEF's solution-focused work will enable the flow of new and more affordable capital into clean energy sectors. These solutions will be designed to address specific market risks that block capital flows. These will include designing, implementation support, and evaluation of policy instruments, insurance products, and incubation funds.

CEEW-CEF was launched in July 2019 in the presence of H.E. Mr Dharmendra Pradhan and H.E. Dr Fatih Birol at Energy Horizons.

cef.ceew.in

The authors



Rishabh Jain
rishabh.jain@ceew.in | @rjain88

Rishabh, a public policy practitioner, has a keen interest in renewable energy. He leads the Market Intelligence vertical at the CEEW Centre for Energy Finance and works closely with key decision-makers in the Indian RE sector. Before joining The Council, he has worked with Vikram Solar, Energo Labs, and Oxford Policy Management. Rishabh completed his master's in public policy from Lee Kuan Yew School of Public Policy, National University of Singapore, and bachelor's in mechanical engineering from Birla Institute of Technology, Mesra, India.

"Cost should no longer be the centrepiece of new installations and our approach needs to be multi-dimensional, which takes into account the country's growth, creation of new jobs, and energy security. This report recommends both short- and long-term interventions to meet domestic needs and become export competitive."



Arjun Dutt
arjun.dutt@ceew.in I @ArjunDutt_00

Arjun Dutt is an Associate at CEEW-CEF. His work is geared towards enhancing the flow of affordable finance towards clean energy in emerging economies. This includes analysing the risks constraining the flow of capital towards clean energy and developing suitable interventions to de-risk investments. Prior to his association with CEEW, Arjun worked for over three years in equity research. He has a bachelor's in electronics and communication engineering from the Delhi College of Engineering and an MBA from the Management Development Institute, Gurgaon.

"Policymaking to support the scaling up of domestic PV manufacturing must strike a fine balance between predictability and dynamism. Predictability is needed in terms of demand-side drivers and measures of policy support. At the same time, policymaking should be nimble enough to respond to emerging developments in the dynamic clean energy space."



Kanika Chawla
kanika.chawla@ceew.in | @kanikachawla8

Kanika is a policy specialist working at the intersection of clean energy and financial markets. She is the Director of the CEEW Centre on Energy Finance and also manages The Council's research and outreach in renewable energy policy, regulation, markets, and socio-economic value. She is actively engaged with private and public enterprises within and outside India in designing and developing financial de-risking instruments.

"To truly develop a thriving solar manufacturing sector, India needs a strategic green industrial policy that allows maximum domestic value capture from the ongoing energy transition. A good policy needs to based on evidence, set clear targets, display long-term vision, and reward innovation."



In addition to solar cells and modules, tariff barriers are currently levied on EVA and solar glass too.

Contents

Executive summary 1

1. Introduction 4

2. Did we learn from the past? 5

3. Key drivers to enhance the competitiveness of domestic PV module manufacturing (from solar cells) 6

Methodology 6

Key assumptions 7

Results 7

4. Recommendations 10

Interventions required in the immediate to short term 11

Interventions required in medium to long term 14

Some interventions only yield short-term benefits: tread with caution 15

5. Conclusion 16

References 17

Annexures 18

Annexure I: Impact matrix of the key interventions in the solar manufacturing sector 18

Annexure II: Key respondents to primary survey 18

Image: Vikram Solar ltd

Figures

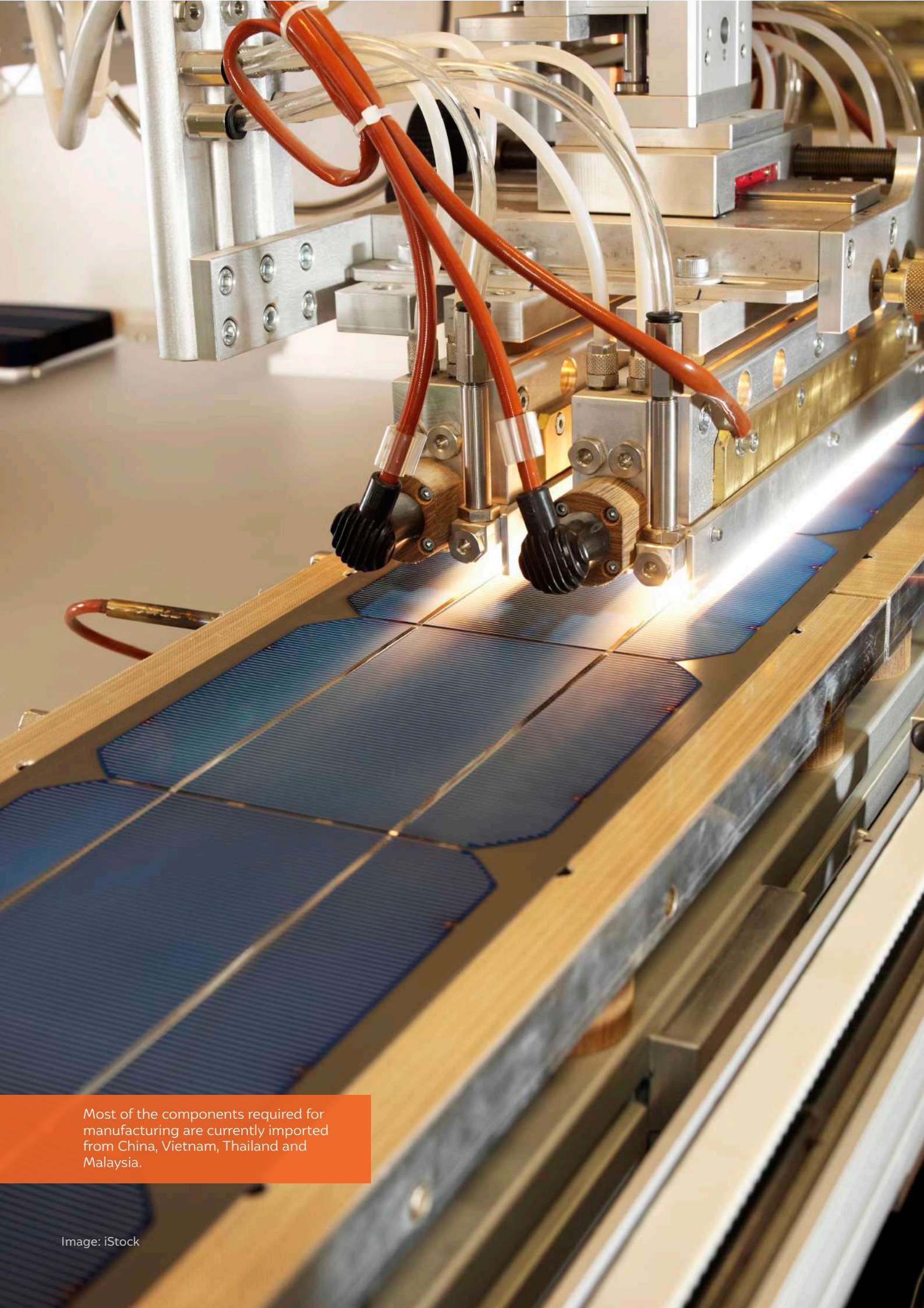
Figure ES1: The difference in the contribution of bill of materials to module selling prices is the highest between India and China 1
Figure 1: Bill of materials constitute the lion's share of Indian module selling price (capacity utilisation: 50%) 7
Figure 2: Electricity, land lease, other overheads, cost of debt, and return on equity contribute insignificantly to Chinese module selling price (capacity utlisation:100%) 8
Figure 3: The difference in the contribution of bill of materials to module selling prices is the highest between India and China 8
Figure 4: Cell contributes heavily to bill of materials cost of Indian manufacturers 9
Figure 5: Chinese module manufacturers procure non-cell BOM at cheaper rates too 9
Figure 6: All components contribute almost equally to bill of materials cost in India and China 10

Tables

Table ES1: Recommended BCD structure on solar cells and modules 2
Table ES2: Recommended design of production subsidy for solar modules 2
Table ES3: Linking VGF with technological improvements 3
Table 1: Recommended BCD structure on solar cells and modules 11
Table 2: Recommended design of production subsidy for solar modules 12
Table 3: Linking VGF with technological improvements 12
Table 4: Recommended deployment targets for solar cell and module manufacturing 13
Table 5: Recommended deployment targets for polysilicon and wafer manufacturing 14

Acronyms

ADD	anti-dumping duty
BCD	basic customs duty
BEL	Bharat Electronics Limited
BHEL	Bharat Heavy Electricals Limited
BOM	bill of materials
CEEW	Council on Energy, Environment and Water
CEEW-CEF	Centre for Energy Finance at the Council on Energy, Environment and Water
CPSU	Central Public Sector Undertaking
CUF	capacity utilisation factor
DCR	domestic content requirement
DGAD	Directorate General for Anti-Dumping and Allied Duties
DGS	Directorate General of Safeguards
DGTR	Directorate General of Trade Remedies
DTA	domestic tariff area
EVA	ethyl vinyl acetate
FiT	feed-in tariff
GBI	Generation Based Incentive
GW	gigawatt
HJT	heterojunction technology
IISc	Indian Institute of Science
IIT	Indian Institute of Technology
INR	Indian Rupees
IREDA	Indian Renewable Energy Development Agency
ISA	International Solar Alliance
M-SIPS	Modified Special Incentive Package Scheme
MeitY	Ministry of Electronics and Information Technology
MNRE	Ministry of New and Renewable Energy
MoC&I	Ministry of Commerce and Industry
MoF	Ministry of Finance
MSME	micro, small, and medium enterprises
MW	megawatt
PERC	passivated emitter and rear cell
PPA	power purchase agreement
PSUs	public sector undertakings
PV	photovoltaic
PVDF	polyvinylidene fluoride
RE	renewable energy
SECI	Solar Energy Corporation of India
SEZ	special economic zone
SGD	safeguard duty
SIPS	Special Incentive Package Scheme
TPT	Tedlar polyester Tedlar
VGF	viability gap funding
Wp	watt peak
WTO	World Trade Organization



Most of the components required for manufacturing are currently imported from China, Vietnam, Thailand and Malaysia.

Image: iStock

Executive summary

India has made rapid advances in the energy transition in recent years, propelled by large-scale additions to solar and wind energy capacity. But the domestic manufacture of solar photovoltaic (PV) cells and modules has not kept pace with this leap in capacity. Still, China accounts for the majority of India's solar PV imports. Certain measures, such as a two-year safeguard duty (SGD) imposed in July 2018 (now extended till July 2021) or the recent award of large manufacturing-linked tenders by the Solar Energy Corporation of India (SECI), to support the domestic manufacturing sector have not yielded much, as the sector is very small to pose any serious competitive challenge.

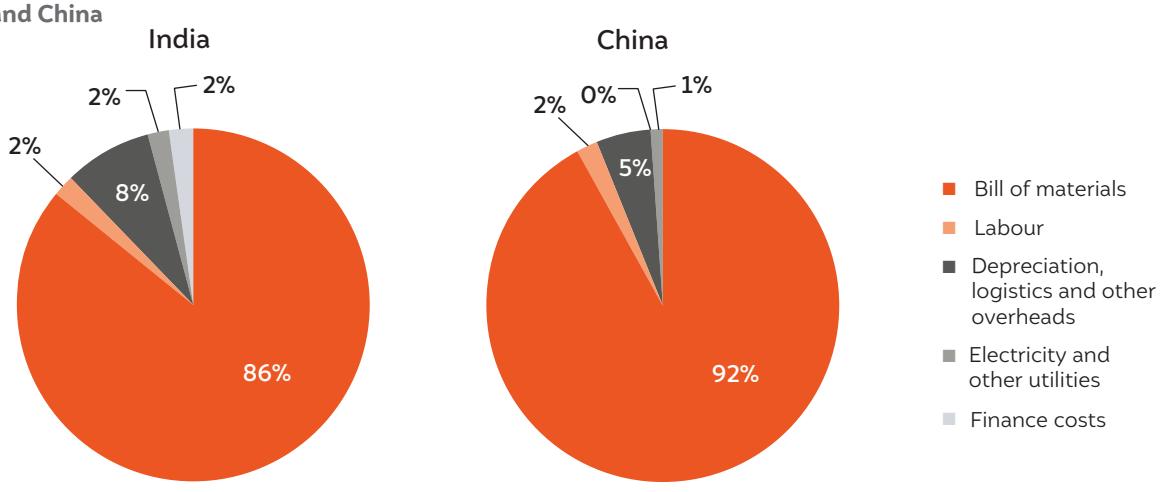
Domestic manufacturing is being pursued with renewed vigour for two reasons. First, Prime Minister Modi gave a clarion call for self-reliance (*Atma Nirbhar Bharat Abhiyaan*) to drive the economic recovery from COVID-19. Second, the tense events at the India–China border have suddenly infused uncertainties into a highly import-dependent sector. The need for enhancing the country's energy security and the ambitious renewable energy targets set for 2022 and 2030 ensure a continuous domestic demand pipeline. Even then, domestic manufacturers are not in a position to gain from this phenomenal opportunity. What ails them and what can be done? To understand the scenario, we empirically assess the competitive advantage that China has over India in solar cell to solar module manufacturing. We then present a suite of interventions that need to be implemented in a comprehensive and balanced manner to yield both short-term relief to the sector, while also making the sector competitive in the long term.

Indian domestic module manufacturing suffers from twin problems of higher pricing and underutilisation of capacity. Indian modules are nearly 33 per cent more expensive than their Chinese counterparts, assuming 50 per cent and 100 per cent capacity utilisation for Indian and Chinese manufacturers, respectively. Capacity utilisation at present in India hovers around 50 per cent as per data gathered from the manufacturers. Indian manufacturers stand to benefit merely by pushing up productivity to full capacity: the resulting cost savings would lower the price differential with China to 22 per cent, thus giving a competitive push.

Figure ES1 provides a snapshot of the components of module selling prices in India and China. The share of bill of materials (BOM) for Indian manufacturers is 86 per cent whereas, other direct costs (labour), overheads (electricity, other utilities, land, logistics, depreciation) and finance costs account for only around 14 per cent of the selling price. The smaller contribution of other costs to unit prices in China is a result of higher scale and superior terms of financing.

An in-depth analysis of selling price of modules using a component-wise contribution paints a different picture. BOM cost accounts for 56 per cent of the difference in selling prices (INR 3.32/watt peak (Wp)) between Indian and Chinese modules assuming same utilisation levels. Direct, overhead, and finance costs contribute another 36 per cent to the difference. The larger scale of manufacturing of Chinese module manufacturers lowers their unit labour and overhead costs. Chinese manufacturers enjoy favourable financing options, which makes an eight per cent contribution to the difference in selling prices.

Figure ES1 The difference in the contribution of bill of materials to module selling prices is the highest between India and China



*Source: CEEW-CEF analysis**
** based on a primary survey conducted with key manufacturers in May 2020. Respondent details in Annexure II*

Our proposed recommendations to enhance India's competitiveness in solar manufacturing touch upon policy responses, a long-term vision, infrastructure development, financial subsidy, and non-financial concessions. Policy or fiscal interventions should aim at (i) an immediate price levelling and (ii) providing systemic solutions to build the competitiveness of the domestic manufacturing sector in the medium run. Several interventions necessitate a balanced implementation to lend support to the sector in a progressive, market-making manner. Our recommendations are divided into three components: (i) short-term interventions to support existing manufacturers survive the current crisis; (ii) interventions for enhancing competitiveness and growth of the sector in the medium to long term; and (iii) the short-term interventions to avoid such that the domestic manufacturing sector does not turn into a white elephant.

Short-term response measures

Our recommendations of short-term measures by the government focus on removing ambiguities in policy and structuring of duties and concessions that boost the domestic manufacturing of solar equipment and components.

Policy uncertainty increases risk perception and delays much needed investments in manufacturing. The government has recently extended the SGD duty but it must provide visibility on the implementation timelines and tenure of the proposed basic customs duty (BCD). See Table ES1 for proposed recommendations.

Table ES1 Recommended BCD structure on solar cells and modules

Timelines	BCD on cells	BCD on modules
Till March 2021	Nil	10%
April 2021 to March 2022	Nil	20%
April 2022 to March 2023	10%	20%
April 2023 to March 2027	20%	20%
April 2027 to March 2030	10%	10%
Beyond April 2030	Nil	Nil

Source: CEEW-CEF analysis

Note: The maximum rate of 20 per cent is in line with that announced in the annual budget. Ministry of Finance may revise the numbers in consultation with the Ministry of Renewable Energy (MNRE)

It is expected that a 20 per cent BCD may lead to an increased tariff of INR 0.20 to 0.25 per unit of electricity. Introducing a tariff barrier alone would not guarantee domestic procurement in the short term. Our analysis suggests that non-hardware cost differential between an Indian and a Chinese manufacturer could vary between INR 1.45/watt and INR 3.20/watt.[1] These include the costs of labour, depreciation, finance, utilities, land, and other overhead expenses. Further, domestic module manufacturers are still dependent on imported cells. If fears of foreign cell manufacturers resorting to artificial increase in prices become a reality, Indian module manufacturers would be burdened further. So we recommend a short-term production subsidy for manufacture of modules using domestic cells to reduce the cost differential between domestic and imported modules (Table ES2).

Table ES2 Recommended design of production subsidy for solar modules

Parameter	Recommendations
Timelines September 2020 to March 2021	**Modules based on domestic cells—INR 2.5/watt Modules based on imported cells—INR 1.5/watt**
Maximum support per manufacturer	250 MW (domestic or imported cells)
Minimum threshold efficiency	19%
Project eligibility	Only utility-scale projects with tariffs discovered via reverse bidding with commissioning dates before March 2021.
Who applies for the subsidy	Module manufacturers to apply for subsidy on behalf of the buyer

Source: CEEW-CEF analysis

We recommend an additional subsidy of INR 1/watt for modules manufactured using domestic cells only for a short term—till March 2021. Under this scheme, by restricting the eligibility of the subsidy to 2.5 to 3 GW modules, the central government is likely to incur a financial burden of nearly INR 600 crore.

We also call for the introduction of domestic procurement programmes such as the *Central Public Sector Undertaking Scheme* (*CPSU*) to mandate the usage of domestic components. In the current CPSU scheme, any government producer can bid and seek a maximum viability gap funding (VGF) of INR 70/watt (MNRE

1 Higher capacity utilisation levels can directly improve the competitiveness of Indian module manufacturers. Further, economies of scale can boost the competitiveness of both domestic BOM components as well as modules. We see that at 100 per cent capacity utilisation, the differential between Chinese- and Indian-manufactured modules is INR 1.45/watt whereas at 50 per cent capacity utilisation the differential is INR 3.20/watt.

2020a). Under this scheme, the sale of power from one government entity to another has been capped at INR 2.8/unit. Manufacturers have in the past suggested that a 0.6 per cent increase in efficiency can reduce the final module costs by up to 3 to 4 per cent. Reducing the tariff to INR 2.5 and allocating projects with a capacity of 2.5 GW to high-efficiency modules would ensure that India can become a global leader in solar module technology. Intense research and development (R&D) efforts should be undertaken to reduce the cost of modules and cells, leading to reduced requirement for VGF (Table ES3).

Table ES3 Linking VGF with technological improvements

Deployment year (2.5 GW/year)	Module efficiency
2021–22	20%
2022–23	21%
2023–24	22%
2024–25	23%

Source: CEEW-CEF analysis

Note: The above numbers are indicative, and the eligibility may differ based on module technologies such as multi/mono crystalline and thin film modules

To stimulate investments in solar manufacturing, the government needs to extend the support in form of policies and concessions to ensure that investors (foreign and domestic) develop an interest in the sector. The Ministry of New and Renewable Energy (MNRE) should set a national target of achieving at least 50 GW of manufacturing capacity by 2030.[2] Yearly addition of 5–10 GW of solar module and cell manufacturing will ensure that new capacity is able to offset the impact of dysfunctional factories.

Currently, exemption from customs duty on importing manufacturing equipment is provided to manufacturers (CBIC 2017) and the MNRE is working towards extending the same benefit for wafer, ingot, and polysilicon manufacturing (MNRE 2020b). The government's policy should be directed towards reducing India's over-reliance on imported manufacturing equipment. One way to incentivise domestic wafer manufacturers or independent technological providers to develop, commercialise, and produce manufacturing equipment indigenously is to award them with cash prizes on achieving specific targets—sales, cost, or innovation (patent filing). The prizes need to be reviewed every year to ensure global competitiveness.

In the Union Budget 2020, the government finally split the HS code used to record the import of solar cells and modules in India. Until now, the HS code 8541 4011 was used for both cells and modules. Now an additional code 8541 4012 has been created, which will record data only for modules whereas the former code will be used for cells. In addition to creating this division, the ministry also needs to record data in capacity terms and not just their monetary value. Currently, there is no way to accurately estimate imports based on government records. To ensure targeted policy design, data in INR and kW units need to be gathered. This will increase both transparency and accountability in the sector.

Systemic market-making interventions

Systemic market-driven interventions are required which are aimed at creating a domestic infrastructure for solar equipment manufacture. There needs to be a thrust on micro, small, and medium enterprises (MSME) for sourcing materials, and indigenous R&D efforts for developing novel technologies in the solar sector.

Removing our reliance on imported raw materials would give impetus to energy security of the country. One option is to look for materials that can be sourced within the country. Quartz is abundantly available in India and, needless to say, the government must incentivise companies that come forward to manufacture solar modules using quartz. BCD may be implemented on solar wafers from 2027 to 2030 (in line with BCD on solar cells and modules). These upstream PV manufacturing stages are capital- and energy-intensive and the government should explore avenues to accelerate their scaling up.

Indian manufacturers rely also on imports for sourcing components. Currently, not just cells but other non-cell BOM components like glass, ribbon, ethylene vinyl acetate (EVA) sheet, and others, which constitute 30 to 35 per cent of the total module cost, are imported as well. Micro, small, and medium enterprises (MSMEs) are producing these components in India, but their quality and price are not competitive. Possible ways to procure BOM components from Indian MSMEs by addressing their shortcomings will go a long way in boosting the MSME sector, create additional jobs, and bring down reliance on China.

A solar power plant does not have many moving parts, thus rendering manufacturing of key equipment, such

2 It has been estimated that the annual global PV installations may increase to 270 GW in 2030 from the current levels of 100 GW/year (IRENA 2019). Also, the global production of solar modules was in the range of 130–140 GW in 2019 (Fraunhofer ISE 2020). If India is able to ramp up its capacity, it will not only be able to service the domestic market but export to other countries too.

as solar modules, inverters, and module mounting structure, easy and scalable. Indian solar manufacturers import most of the components and raw materials used in the manufacture of key equipment. We therefore recommend setting up of 10 GW solar manufacturing parks near the ports in the states of Maharashtra, West Bengal, Andhra Pradesh, and Tamil Nadu to boost domestic manufacture of solar plant equipment. The government also, in parallel, should ensure availability of suitable infrastructure in these locations. To provide further boost, tax holiday and duty waiver could be offered to these manufacturers. The states should be entrusted with the responsibility of bringing in the skilled workforce to build self-reliance.

European countries have succeeded in not only identifying new technologies like organic solar modules at the lab scale through consistent R&D effort, but also in deploying them. India's prestigious technology institutions such as the Indian Institute of Technology (IIT) Bombay, IIT Delhi, IIT Madras, and the Indian Institute of Science (IISc) Bangalore could be roped in, through specific financial grants, by the government to develop novel solar technologies and also commercialise them to make India self-reliant.

Interventions to avoid

We also feel it is imperative to have a broad vision and a long-term perspective for the solar manufacturing sector in India. Interventions providing short-term benefits may be either unsustainable in the medium to long run or do little to build competitive advantage for the sector. A long-term commitment to developing local manufacturing should be pledged by the government through support mechanisms. Manufacturers must be encouraged and rewarded for investing in R&D and technology improvements. The vision of domestic solar manufacturing eventually maturing to serve a global market should be pursued rather than just developing a captive protected domestic market. To realise this vision, all support interventions must be accompanied by clearly defined sunset clauses.

Policymakers and financiers cannot be expected to be aware of the rapid scale of global technological advancements. We thus call upon the Indian Renewable Energy Development Agency (IREDA) to take a lead role in providing technical expertise to the government, banks, and lending institutions on solar manufacturing to avoid the government favouring obsolete technologies.

With favourable policies, RE manufacturers, like RE developers, will also be able to raise low-cost capital to scale up their facilities in the medium run. Hence, the government's endeavour should be aimed at developing a vibrant solar manufacturing ecosystem to encourage innovative entrepreneurs to pursue solar manufacturing. Merely offering project finance is not likely to fetch long-term dividends. Providing visibility of a demand pipeline and support mechanisms that use limited pools of public money more effectively may have a better impact on solar manufacturing sector rather than costly project financing with public funds.

Finally, lessons from several sectors suggest that the government must focus on interventions that can be provided in a time-bound manner. Further, overlapping mandates of different ministries like the MNRE, Ministry of Commerce and Industries (MoC&I), Ministry of Finance (MoF), and Ministry of Electronics and Information Technology (MeitY) should be ironed out and a mechanism developed to coalesce their efforts to lend efficiency and create impact.

The government should not view support for the solar manufacturing sector as an isolated policy decision. For creating lasting impact, various concerns need to be addressed, different ministries should work in tandem, and these efforts should be balanced by nimble individual interventions from time to time. We provide evidence for an imbalanced playing field for manufacturers at present and propose a toolbox of policy levers that the government must deploy to create a globally competitive solar manufacturing sector in India by 2030.

1. Introduction

Indian solar manufacturers have not been able to exploit India's ambitious energy transition, as they have not kept pace with the global advancements in technology. The weak link in the grand scale of RE plans is the country's high dependence on imports for solar modules and the raw materials required for its manufacturing. Globally, individual companies have capacities that are greater than the total sum of India's manufacturing capacity (10 GW solar module and 3 GW solar cell) (MNRE 2020c). The COVID-19 pandemic sweeping the world at present and recent geo-political developments have raised concerns about the country's energy security. But unfortunately, there is no silver bullet that can solve this problem. We provide a brief context, investigate the reasons for the higher cost of Indian solar solar modules, and recommend some key interventions that will be required to scale up domestic manufacturing of crystalline solar photovoltaic (PV) modules in the country.

2. Did we learn from the past?

A critical observation of two key features of India's energy transition shows that focus is more on clean energy deployment than on energy security. Indian RE strategy is driven by optimising for tariff decline of renewable power and ad-hoc policy support for solar manufacturing. Many countries rely on a feed-in-tariff (FiT) regime to scale up RE projects (until recently, even China and the United States have opted for FiT regimes). Instead, India resorted to reverse bidding right from the early stages of the *National Solar Mission*. This aggressive tariff decline became the most significant marker of India's renewable energy scale-up, consequently detrimental to domestic industry. Competitive unit economics drove developers to prefer low-cost imports over relatively expensive domestic products. Domestic cell and module manufacturers, already on the backfoot, neither have the scale, time, nor financial resources to catch up to global competition.[3]

Inconsistent policy support for domestic manufacturing of PV cells and modules put domestic manufacturers at a further disadvantage. From designing the domestic content requirement (DCR) in 2011 to promising basic customs duty (BCD) in 2020, the Indian government promised much to domestic manufacturers but ended up delivering little. In the early stages of the National Solar Mission, the government implemented the DCR policy giving preference to domestically produced polycrystalline PV modules and provided financial support to such projects. However, exemption of thin film modules distorted the market. This exemption led to increased deployment of thin film modules in solar projects due to their price advantage, and the ratio of thin film to polycrystalline deployment in India was inverted compared to the global average. The policy was contested by the United States and did not stand the scrutiny of global trade rules in the courts of the World Trade Organization (WTO), and India had to withdraw the scheme (WTO 2016). In anticipation of increased orders under the DCR policy, many manufacturers had set up factories and applied for capex support under the *Special Incentive Package Scheme* (*SIPS*) and *Modified Special Incentive Package Scheme* (*M-SIPS*).[4] The delay in the disbursal of 20 to 25 per cent capex support under these two schemes widened the gap with respect to imported products. Few manufacturers even approached the Directorate General for Anti-Dumping and Allied Duties (DGAD) and the Directorate General of Safeguards (DGS) to impose tariff barriers on imported products. In 2014, while the then DGAD recommended the implementation of anti-dumping duty, the Ministry of Finance rejected the proposal and instead promised greater off-take for domestically produced modules under the *CPSU* scheme. The promise was unkept for years, and manufacturers this time approached the DGS again to impose a safeguard duty.

In 2018, the Ministry of Finance imposed a safeguard duty (SGD) at short notice, creating some avoidable upheaval and confusion (on applicability) in the sector. The duty was imposed for two years, too short a time frame to have any impact on supporting domestic

> **The safeguard duty was imposed for two years, too short a time frame to have any impact on supporting domestic manufacturing.**

3 The Indian wind sector enjoyed support in form of FiT and generation-based incentive (GBI) for a long time (15 years) until the sector moved to reverse bidding in 2017. FiT and GBI, coupled with accelerated depreciation, not only increased project deployment but supported domestic manufacturing industry, which in turn relied on thousands of micro, small and medium enterprises (MSMEs) to supply key components and spare parts. Currently turbines deployed in India have very high degree of indigenisation (70–90 per cent).

4 The Ministry of Electronics and Information Technology (MeitY) through the *Special Incentive Package Scheme (SIPS) and Modified Special Incentive Package Scheme (M-SIPS)* promised both capex support and production subsidy to solar module, cell, and wafer manufacturers in special economic zones (SEZ) and domestic tariff area (DTA). However, even with in-principle approvals, many manufactures have been anticipating disbursal for many years.

manufacturing. Investors too felt two years was too short a window to set up a new manufacturing facility. Meanwhile, the project developers continued with imports, hoping compensation under the *change in law* clause of the contract would cover the changed terms. This eventually proved to be counter-productive for manufacturers since imported products became cheaper when compared to domestically manufactured products. The SGD did not seem to matter any longer. For new bids, manufacturers preferred import modules as the SGD would expire sooner, or its 15 or 20 per cent range still made imported modules more competitive than domestically manufactured modules. Recently on 29th July 2020, the government extended the safeguard duty on solar cells and modules by one year. A 14.9 per cent duty would be applicable on imports from 30th July 2020 to 29th January 2021 and a 14.5 per cent duty would be applicable on imports from 30th January 2021 to 29th July 2021. The order is applicable on imports from all developed countries, China, Vietnam and Thailand (MoF 2020). Unlike the 2018, safeguard duty order, Malaysisa has been exempt which opens the door for increased imports from the country. In February 2020, the government announced its intent to impose BCD on solar cells and modules, which added to the existing activity in the sector.

To ensure reliability of solar PV cells and modules, the MNRE had also issued an order in 2019 to enlist solar PV models and manufacturers to be used in all projects that are bid out as per the central government standard bidding guidelines and are either government-owned or government-assisted. However, this order has also been extended to September 2020 from the earlier implementation date of 1 April 2020 due to the prevailing COVID-19 pandemic (MNRE 2020). This move by MNRE was also indirectly aimed at avoiding the inflow of low-quality imported products into India. It is also important to note here that in December 2017, the MNRE shared a draft policy document to scale up domestic manufacturing, with a target to set up 10 GW of manufacturing over five years. The domestic manufacturers are still hopeful this policy would change their fortunes.

This sets the context for a further discussion because incoherent planning and messaging have put domestic solar module and cell manufacturers at a greater disadvantage and clamped their competitiveness. To ensure the successful implementation of any policy, the government needs to gain the trust of key stakeholders and provide them a commitment of support—regulatory and financial—which will ensure survival of those who are the intended beneficiaries in the short term and increase their competitiveness in the long run.

> To ensure the successful implementation of any policy, the government needs to gain the trust of key stakeholders.

3. Key drivers to enhance the competitiveness of domestic PV module manufacturing (from solar cells)

There are four key steps in the manufacture of solar modules—quartz to polysilicon; polysilicon to wafer; wafer to cell, and finally cell to solar module. As mentioned earlier, India currently has a manufacturing capacity of 10 GW of solar modules (from solar cells) and 3 GW of solar cells (from wafers). In May 2020, we reached out to domestic module manufacturers with annual manufacturing capacity ranging from 100 MW to 2,000 MW. A detailed questionnaire, with the intent to analyse the breakup of manufacturing cost in India, was shared with them. In addition, based on their experience, the manufacturers were asked to share similar details for China. Using the inputs from the manufacturers, a detailed analysis was conducted. **However, we note that the comparative analysis that is provided below is limited to manufacturing of solar modules from solar cells.**

A number of underlying determinants are used to arrive at the selling prices of solar PV modules—including costs and returns to debt and equity investors. The values of these determinants vary across markets, which manifest as differences in selling prices. When comparing the prices of the solar modules produced in India and those imported from China, the country's largest source of solar module imports, the variations in various determinant values is significant. Segmenting the module price into its determinants would aid in ascertaining why Indian modules are at a competitive disadvantage. This, in turn, facilitates an assessment of the impact of targeted measures geared towards improving the competitiveness of domestic module manufacturing.

Methodology

The unit selling price of a module (in INR/Wp) covers all direct and indirect costs involved in the process of

production as well as investor returns. Aggregating these determinants applicable to unit module capacity makes up the unit selling price. Direct costs such as raw materials and the variable components of labour and electricity costs can be directly allocated to unit module capacity. In addition, overheads such as other utilities, land lease expenses, and logistics costs were estimated for the entire plant and then allocated to unit module capacity based on an assumed capacity utilisation factor (CUF). Returns for debt and equity investors were also estimated at the plant level and allocated to unit module capacity considering an assumed capacity utilisation level.

Data on these cost drivers for both Indian and Chinese modules was sourced from domestic module manufacturers (names of the manufacturers are provided in Annexure II). The following analysis is based on the mid-points of data inputs provided by these manufacturers for both Indian and Chinese module manufacturing.

Key assumptions

- Production capacity (mono passivated emitted and rear cell (PERC) modules, assumed average size of facilities):
  - Indian module manufacturing plant capacity: 500 MW
  - Chinese module manufacturing plant capacity: 2,000 MW
- Plant's capital expenditure (solar cell to solar module):
  - Indian manufacturing (INR crore/MW): 0.3
  - Chinese manufacturing (INR crore/MW): 0.2
- Plant's useful life:
  - Indian manufacturing: 5 years
  - Chinese manufacturing: 5 years
- Capacity utilisation:
  - Indian manufacturing: 50%[5]
  - Chinese manufacturing: 100%
- Return on equity (pre-tax):
  - Indian manufacturing: 18%
  - Chinese manufacturing: 10%

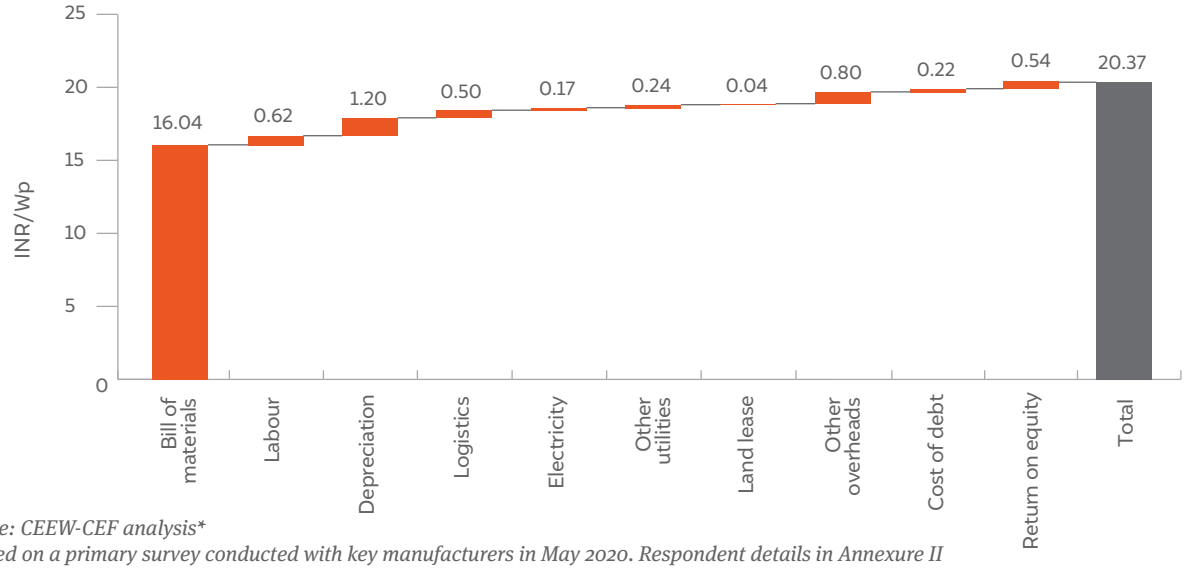
Results

To understand the magnitude of difference in prices, Figures 1 and 2 present a breakdown of Indian and Chinese module prices into their constituents. Based on our analysis, Indian modules are nearly 33 per cent more expensive than their Chinese counterparts.

In our analysis, we have respectively assumed 50 per cent and 100 per cent capacity utilisation for Indian and Chinese manufacturers. We stress that full utilisation of capacity can significantly impact the final cost and increase competitiveness of the domestic players.

In order to understand the fundamental differences in the competitiveness of Indian and Chinese module manufacturers, we need to first control for the distortionary effect of differences in capacity utilisation. If Indian module manufacturers were to operate at 100 per cent utilisation levels, the difference in prices would narrow to INR 3.32/kWh or around 22 per cent. The analysis presented below assumes both Indian and Chinese module manufacturers operating at full capacity utilisation levels.

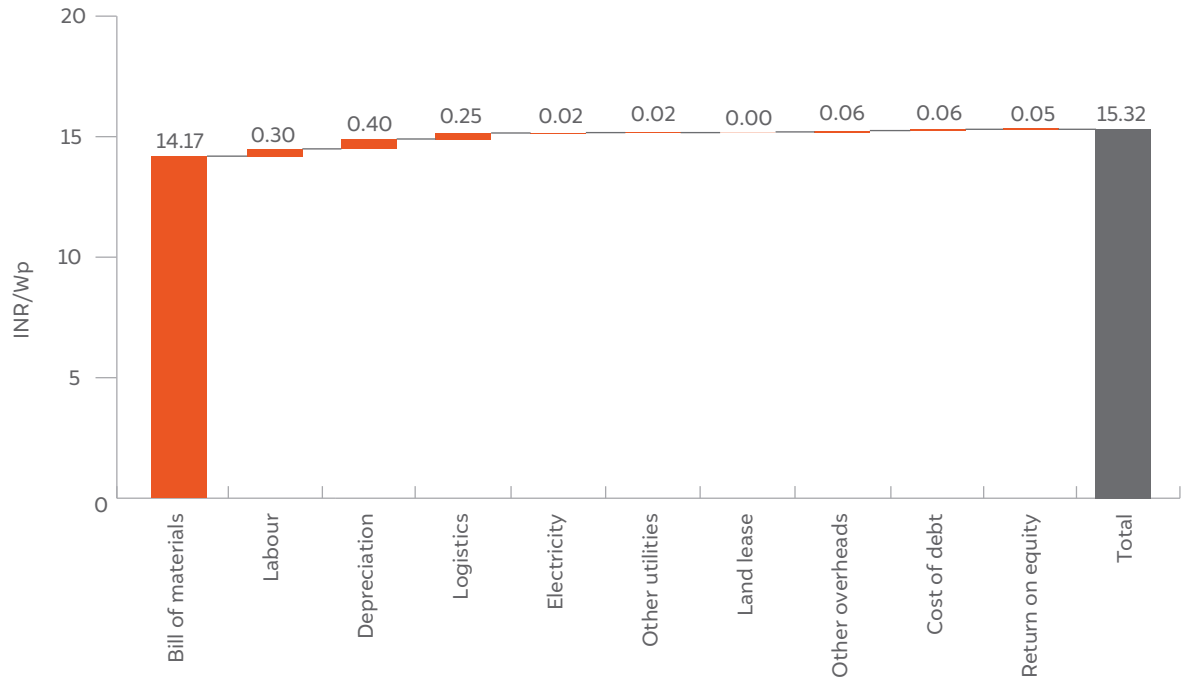
Figure 1 Bill of materials constitute the lion's share of Indian module selling price (capacity utilisation: 50%)



*Source: CEEW-CEF analysis**
** based on a primary survey conducted with key manufacturers in May 2020. Respondent details in Annexure II*

5 Assumed average based on our consultations with the industry.

Figure 2 **Electricity, land lease, other overheads, cost of debt, and return on equity contribute insignificantly to Chinese module selling price (capacity utlisation:100%)**



*Source: CEEW-CEF analysis**

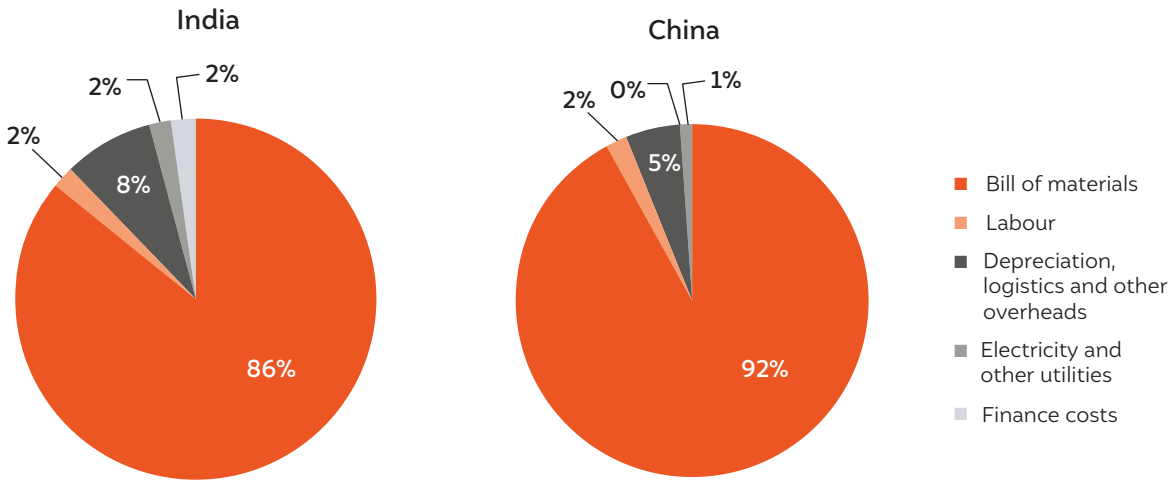
Major determinants of Indian module selling prices

- The bill of materials (BOM) (disaggregated further in Figures 4 and 5) accounts for nearly 86 per cent of the selling price in India (Figure 3). Other direct costs (labour), overheads (electricity, other utilities, land, logistics, depreciation), and finance costs account for around 14 per cent of the selling price. Similarly, BOM cost determines 92 per cent of module selling price in China, while other costs total up to 8 per cent of the unit prices. This is attributed to higher scale and superior terms of financing in China.

Reasons for differences in module prices

- Around 56 per cent of the difference in selling prices (INR 3.32/Wp) between Indian and Chinese modules at the same utilisation levels arises from differences in the BOM costs (Figures 2).

Figure 3 **The difference in the contribution of bill of materials to module selling prices is the highest between India and China**

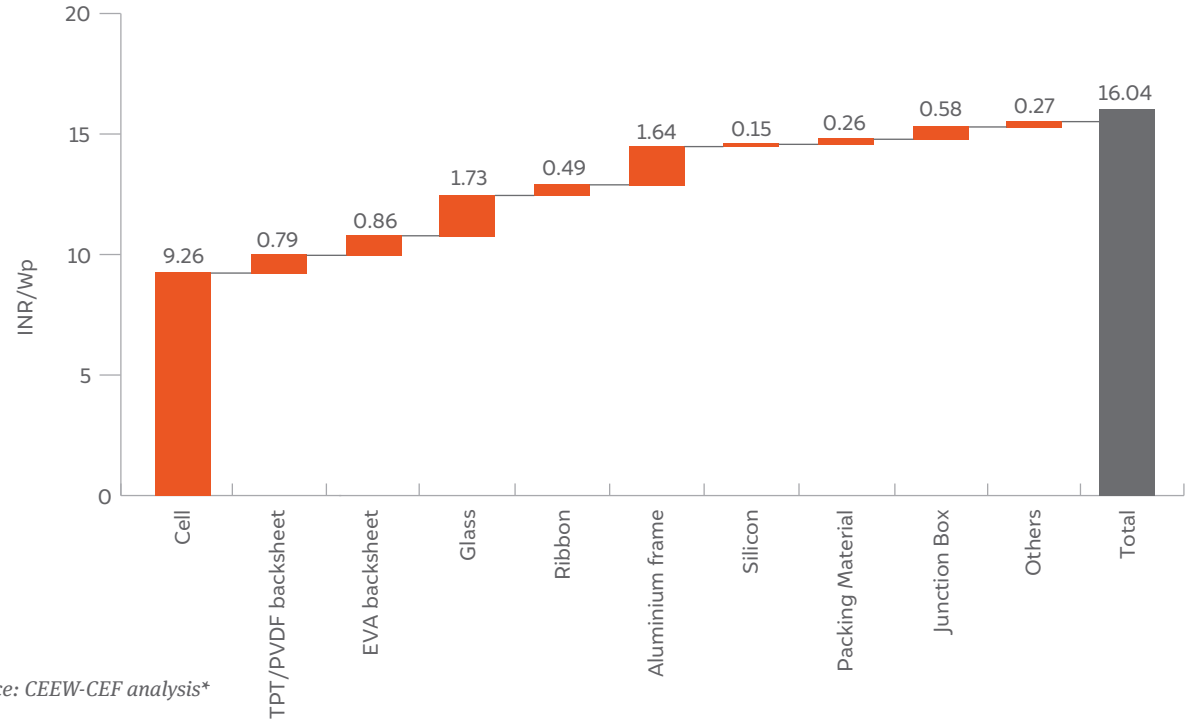


*Source: CEEW-CEF analysis**
** based on a primary survey conducted with key manufacturers in May 2020. Respondent details in Annexure II*

- The remainder of the difference may be explained by differences in other direct, overhead and finance costs. The larger scale of manufacturing of Chinese module manufacturers translates into lower unit labour and overhead costs and accounts for another 36 per cent of the difference in selling prices. Further, lower financing costs for Chinese manufacturers explain another 8 per cent of the difference in selling prices.

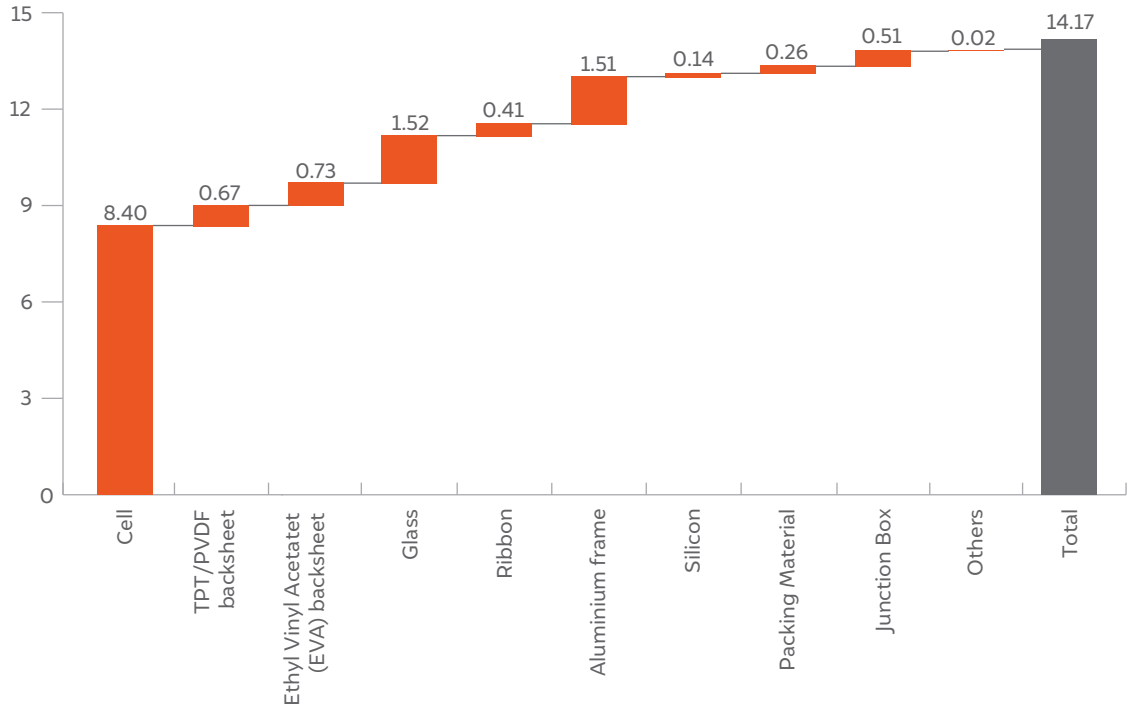
Since the BOM is the key differentiator, Figures 4 and 5 present a comparison between BOM costs for Indian and Chinese manufacturers.

Figure 4 Cell contributes heavily to bill of materials cost of Indian manufacturers



*Source: CEEW-CEF analysis**

Figure 5 Chinese module manufacturers procure non-cell BOM at cheaper rates too

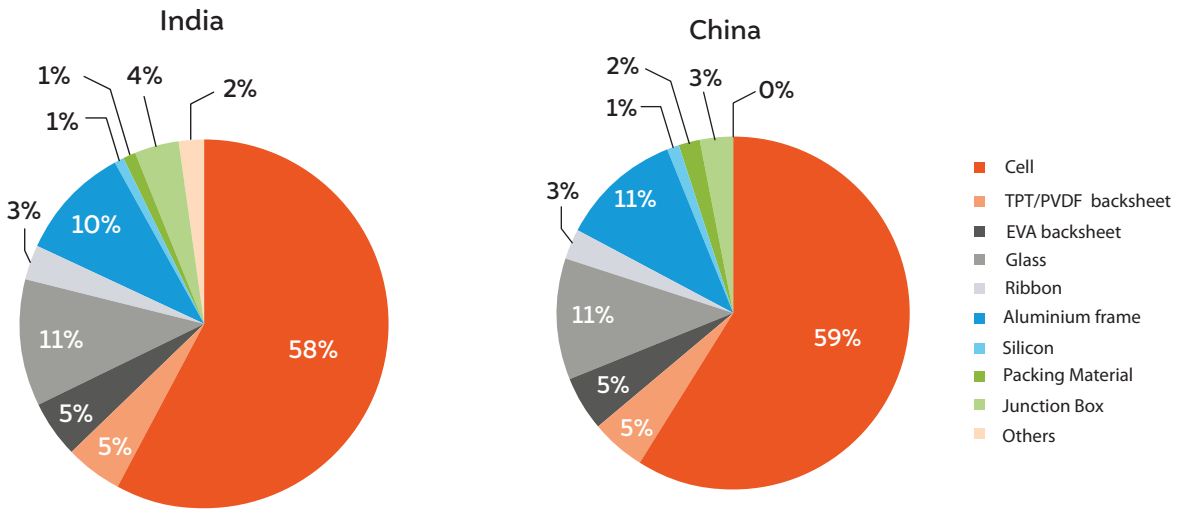


*Source: CEEW-CEF analysis**

** based on a primary survey conducted with key manufacturers in May 2020. Respondent details in Annexure II*

Major determinants of Indian BOM costs

Figure 6 All components contribute almost equally to bill of materials cost in India and China



*Source: CEEW-CEF analysis**
** based on a primary survey conducted with key manufacturers in May 2020. Respondent details in Annexure II*

As is evident from Figure 6, there contribution of components to BOM cost is almost the same in India and China for solar module manufacture. Cells account for around 58 per cent (China, 59 per cent) of Indian BOM costs (Figure 6). The Tedlar polyester Tedlar (TPT)/ polyvinylidene fluoride (PVDF) backsheet, ethyl vinyl acetate (EVA) backsheet, glass, and aluminium frame add up to another 31 per cent of the BOM cost (similar in China).

Differences in BOM costs

- From Figures 5 and 6, nearly 46 per cent of the difference (INR 1.87/Wp) in BOM costs stems from more expensive cells sourced by Indian manufacturers. Differences in TPT/PVDF backsheets, EVA backsheets, glass, and aluminium frame costs account for another 32 per cent of the difference in costs.
- Indian manufacturers rely on imports for most of these raw materials:
  - Solar cells—China, Vietnam, and Thailand
  - TPT/PVDF sheets—Vietnam
  - EVA backsheets—Malaysia (though there is considerable domestic procurement too)
  - Glass—Malaysia
  - Ribbons, aluminium frames, junction boxes—China
- The imposition of duties for some materials (SGD for cells, anti-dumping duty (ADD) on EVA backsheet and glass) further raises the cost. In addition, Indian importers have raised concerns of artificial increase in export prices by foreign suppliers (primarily China).[6]
- The ADD on EVA backsheet and glass is not applicable to module manufacturers in the special economic zone (SEZ). This creates another layer of challenge for manufacturing facilities in domestic tariff area (DTA).

4. Recommendations

India has set a target of solar installed capacity of 100 GW by 2022, which means projects with a capacity of 65 GW have to be installed in the next 30 months. In pursuit of this target, we recommend the key steps that the government needs to undertake to scale up domestic solar manufacturing in India. The recommendations are divided into three components: (i) interventions in the short term, which will avoid the shutdown of factories and help many operational manufacturers to survive the current crisis; (ii) steps to be taken that will have a long-term impact on the sector; and (iii) the pitfall of short-sightedness. A summary matrix of the interventions have also been provided in Annexure I. We classify the recommendations into fiscal and non-fiscal measures. While imposition of duty would bring significant revenue to the government, some measures would require a few ministries to forego revenue. However, the financial implication of the recommended interventions would be limited, and the collection from SGD duty over the past two years (approximately INR 5,000 crore) may be used to cover for these interventions.

6 Based on our consultations with the industry.

Interventions required in the immediate to short term

(a) Clarity on applicability and timelines of tariff barrier

CEEW-CEF recommends the Ministry of Finance (MoF) to implement this intervention

Policy uncertainty increases risk perception and delays much needed investments in manufacturing. The Ministry of Finance has already extended SGD by one year (till July 2021) and now the government needs to provide clear information on the implementation timelines and tenure of the proposed basic customs duty (BCD). This will not only help manufacturers set up factories but also help developers plan their procurement better. The duty must be levied after accounting for construction time and the life of the manufacturing unit (around 10 years). We suggest levying a tapered and a differential BCD on solar cells and modules (Table 1). A higher duty on modules and lower duty on cells would ensure that our module facilities can increase their capacity utilisation factor (CUF) and become competitive with imports. Similarly, a tapered duty on cells has been proposed taking into account the time required for commissioning of new cell manufacturing facilities.

Table 1 Recommended BCD structure on solar cells and modules

Timelines	BCD on cells	BCD on modules
Till March 2021	Nil	10%
April 2021 to March 2022	Nil	20%
April 2022 to March 2023	10%	20%
April 2023 to March 2027	20%	20%
April 2027 to March 2030	10%	10%
Beyond April 2030	Nil	Nil

Source: CEEW-CEF analysis

Note: The maximum rate of 20 per cent is in line with that announced in the annual budget. Ministry of Finance may revise the numbers in consultation with MNRE

At the current levels of tariff and module prices, it is expected that the with 10 and 20 per cent increase in module prices, the tariff may increase by 5 and 10 per cent, respectively. So a 20 per cent BCD may lead to an increased tariff of INR 0.20 to 0.25 per unit of electricity.

However, any decrease (artificial or real) in module prices may compensate for this increase. This trend was witnessed in the last two years when the decline in module prices was steeper than the imposed SGD (meant to deter imports). It should also be noted that project developers will be eligible for reimbursement under the change in law clause for projects that have been bid out before the imposition of the duty. Hence, a clear mechanism needs to be designed so that developers are not burdened with extra cost of procurement. A time-bound reimbursement or exemption for exiting projects should also be planned.

It is, however, important that certain revisions in regulatory provisions are undertaken before the implementation of BCD. Firstly, the Ministry of Finance must devise a mechanism to ensure that manufacturing units in SEZ and DTA are at par for any tariff barriers like basic customs and safeguard duty. Implementing BCD without any changes in customs rules and regulations will be counter productive for domestic solar module manufacturers located in SEZ since BCD will be applicable on the value of the product moved from SEZ to DTA. It needs to be ensured that the duty is not levied on the domestic value add component. Secondly, it should also be noted that solar cells and modules is currently exempted under the Information Technology Agreement-1 (ITA-1). Any move by India to unilaterally impose duties can lead to litigation against India in the WTO. In the past, we have seen Japan and the European Union (EU) initiate dispute proceedings when India imposed customs duty on products exempted under ITA-1 (WTO 2020).

(b) Production subsidy

CEEW-CEF recommends the Ministry of New and Renewable Energy (MNRE) to implement this intervention

A tariff barrier cannot guarantee domestic procurement in a short term because of the change in law clause. Even with significant duty, the current status quo may continue for the next 15–18 months. Our analysis suggests that non-hardware cost differential between an

Indian and a Chinese manufacturer can vary between INR 1.45/watt to INR 3.20/watt.[7] These include the costs of labour, depreciation, finance, utilities, land, and other overhead expenses. In addition, due to limited cell manufacturing capacity, module manufacturers are still dependent on imported cells. In the past, concerns have been raised around an artificial increase in cell prices to reduce the competitiveness of Indian module manufacturers. A sudden surge in demand in the post-COVID-19 scenario can also lead to dumping of low-quality products. A short-term intervention in the form of production subsidy will serve two purposes: (i) reduce cost differential between domestic and imported modules and (ii) ensure deployment of high-efficiency products. We suggest that the domestic module manufacturers be given a production subsidy (Table 2) for modules manufactured till March 2021.

Table 2 Recommended design of production subsidy for solar modules

Timelines—September 2020 to March 2021	Modules based on domestic cells—INR 2.5/watt Modules based on imported cells—INR 1.5/watt
Max support per manufacturer	250 MW (domestic or imported cells)
Minimum threshold efficiency	19%
Project eligibility	Only utility-scale projects with tariffs discovered via reverse bidding with commissioning dates before March 2021
Who applies for the subsidy	Module manufacturers to apply for subsidy on behalf of the buyer

Source: CEEW-CEF analysis

Given that capacity utilisation and economies of scale play a very important role in reducing the cost differential, we suggest limiting the subsidy support to 250 MW per manufacturer. This ensures that small players are able to run at full capacity whereas larger companies with benefits of scale can increase their CUF to become competitive with imported products. An additional support of INR 1/watt has been suggested for modules manufactured with domestic cells. Also, production subsidy should only be applicable only on a short term—till March 2021.

Under this scheme, it is expected that only 2.5–3 GW of modules produced may be eligible for subsidy, which would entail a financial outflow of nearly INR 600 crore for the government. The subsidy may be disbursed to the procurer once the project commissioning certificate is obtained by giving details of the DC and AC capacity of the module. The MNRE may decide to increase the duration and extend support for exports based on availability of budget and in accordance with WTO regulations.

(c) WTO-complaint DCR policies

CEEW-CEF recommends the Ministry of New and Renewable Energy (MNRE), Solar Energy Corporation of India (SECI) and Indian Renewable Energy Development Agency (IREDA) to implement this intervention

India had to change its domestic content requirement (DCR) policies due to the WTO ruling in 2016 (WTO 2016). However, domestic procurement programmes like the *Central Public Sector Undertaking Scheme (CPSU)* can still be designed, which mandate usage of domestic components. Any government producer can bid under this scheme and seek a maximum VGF of INR 70/watt (MNRE 2020a). Under this scheme, the sale of power from one government entity to another has been capped at INR 2.8/unit. The scheme currently lapses in 2023, but the government must present a clear pipeline of projects supporting domestic manufacturing till 2030. In addition, the scheme must also mandate usage of locally produced polysilicon, ingots, and wafer at a later stage in a phased manner.

Manufacturers have in the past suggested that a 0.6 per cent increase in efficiency can reduce the final module costs by up to 3 to 4 per cent. The top runner programme in China helped increase the efficiency of passivated emitted and rear cell (PERC) and N-type technologies. The efficiency of solar modules increased from 22 to 23 per cent in 2019 due to the top runner programme. In the absence of any support, the efficiency would have increased by just 0.3 per cent (Xiao 2020). On the above lines, we recommend reducing the tariff to INR 2.5 and allocating projects with a capacity of 2.5 GW to high-efficiency modules annually, which will ensure that India can become a global leader in solar module technology (Table 3).

7 Higher capacity utilisation levels can translate into considerable improvements in the competitiveness of Indian module manufacturers. Further, economies of scale can boost the competitiveness of both domestic BOM components as well as modules. We see that at 100 per cent capacity utilisation, the differential between Chinese and Indian manufactured modules is INR 1.45/watt whereas with 50 per cent capacity utilisation, the differential is INR 3.20/watt.

Table 3 Linking VGF with technological improvements

Deployment year (2.5 GW/year)	Module efficiency
2021–22	20%
2022–23	21%
2023–24	22%
2024–25	23%

Source: CEEW-CEF analysis

Note: The above numbers are indicative, and the eligibility may differ based on module technologies like multi/mono crystalline and thin film modules

Research and development (R&D) efforts should be directed at lowering the cost of modules and cells, which would lead to a reduced VGF requirement. For encouraging companies to invest in R&D, the government needs to provide a clear and visible schedule of support and demand for domestic manufacturing of PV modules, cells, wafers, and polysilicon. It is also expected that manufacturers would be willing to deploy next generation technologies like heterojunction technology (HJT) if a clear demand pipeline exists.

(d) Commitment and clear timelines

CEEW-CEF recommends the Ministry of New and Renewable Energy (MNRE), Ministry of Commerce and Industry (MoC&I) and Ministry of Electronics and Information Technology (MeitY) to implement this intervention

A continued commitment from the government is very essential for scaling up domestic manufacturing in India. To stimulate investments, the government needs to clearly articulate the schedule and timeline of the various forms of support (policies and subsidies) extended to solar manufacturing. This would encourage investors (foreign and domestic) to respond to the market-support signals from the government and develop an interest in the sector. The MNRE should set a national target of installing at least 50 GW of manufacturing capacity by 2030.[8] Yearly addition of 5–10 GW of solar module and cell manufacturing will ensure that new capacity is able to offset the impact of dysfunctional factories. Broadly, at least 25–30 GW of solar cell and module manufacturing should be based on the latest technology.

Therefore, we suggest setting up of nearly 40 GW and 45 GW of additional module and cell capacity based on advanced technology in the next decade to develop India into a leading hub of PV manufacturing for servicing both domestic and international demand. We recommended a graded target for the government (Table 4) and also call upon the government to ensure a conducive and stable regulatory environment with an assured market for solar manufacturers.

Table 4 Recommended deployment targets for solar cell and module manufacturing

Target year (cumulative capacity)	Wafer to cell (GW)	Cell to module (GW)
FY 2021–22	5	15
FY 2022–23	15	20
FY 2023–24	20	25
FY 2024–25	25	30
FY 2025–26	30	35
FY 2026–27	35	40
FY 2027–28	40	45
FY 2028–29	45	50
FY 2029–30	50	50

Source: CEEW-CEF analysis

(e) Producing manufacturing equipment in India

CEEW-CEF recommends the Ministry of Finance (MoF), Ministry of New and Renewable Energy (MNRE) and Department of Science & Technology (DST) to implement this intervention

Currently, manufacturers of solar cells and modules are exempt from paying customs duty on importing manufacturing equipment (CBIC 2017). The MNRE is also working with the Department of Revenue to update the list of equipment required for manufacturing and extend the benefits to wafer, ingot, and polysilicon manufacturing (MNRE 2020b).

Reducing our reliance on imported manufacturing equipment should be the government's priority. There have been instances in the past when many wafer manufacturers increased the size of wafers, which required adjustments in the cell and module manufacturing facilities.

8 It has been estimated that the annual global PV installations may increase to 270 GW in 2030 from the current levels of 100 GW/year (IRENA 2019). Also, the global production of solar modules was in the range of 130–140 GW in 2019 (Fraunhofer ISE 2020). If India is able to ramp up its capacity, it will not only be able to service the domestic market but export to other countries too.

Wafer size has often varied with the whim of the foreign manufacturer. The standard wafer size has increased from 125 mm x 125 mm in 2010 to 158.75 mm x 158.75 mm in 2019. However, many wafer manufacturers have now started producing wafers of a non-standard size, with some even producing wafer of the size of 210 mm x 210 mm. Lack of standardisation can potentially increase the delivered costs to consumers (Hutchins 2019). However, recently seven leading wafer manufacturers have come to a consensus to establish a new standard size of 182 mm x 182 mm (LONGi 2020). It should also be noted that many of these manufacturers also have control over the technological development of manufacturing equipment too. Any change in their offering upstream would require significant changes in the downstream infrastructure for Indian cell and module manufacturers. To avoid any shock in the future, we need to invest significantly in polysilicon and wafer R&D and equipment manufacturing.

We recommend to the government to institute cash prizes to domestic wafer manufacturers or independent technological providers who can develop, commercialise, and produce manufacturing equipment indigenously. Cash prizes should be linked to meeting of specific targets—sales, cost, or innovation (patent filing). The prizes need to be reviewed every year to ensure global competitiveness. Academic and research institutes need to be called to collaborate with the government to ensure support for new technologies. Technological development may also be achieved by a successful joint venture with foreign companies. or through technology transfer. A similar trend was witnessed in the wind sector in the 1980s and 1990s. The government must focus on country's self-reliance in manufacturing equipment in the next 10 years. India should also incorporate solar manufacturing as a focus industry in its foreign policy and leverage the expertise that exists in Europe, Japan, and the United States of America through collaboration with those governments and private parties.

(f) Better recording of data

CEEW-CEF recommends the Ministry of Commerce and Industry (MoC&I) to implement this intervention.

In the Union Budget 2020, the government finally split the HS code used to record import of solar cells and modules in India. Until now, the HS code 8541 4011 was used for both cells and modules. Now an additional code 8541 4012 has been created, which will record data only for modules whereas the former code will be used for cells. In addition to creating this division, the ministry also needs to record data in capacity terms and not just their monetary value. Currently, there is no way to accurately estimate imports based on government records. To ensure targeted policy design, we need to collect data in INR and kW units.

In addition, to remove biases and improve bankability, the MNRE may regularly release generation data of projects based on domestic and imported modules. This will go a long way in increasing both transparency and accountability in the sector.

Interventions required in medium to long term

(a) Backward integration

CEEW-CEF recommends the Ministry of New and Renewable Energy (MNRE), Ministry of Commerce and Industry (MoC&I) and Ministry of Electronics and Information Technology (MeitY) to implement this intervention

India needs to reduce its reliance on imported raw materials. Quartz is abundantly available in India and incentivising companies that are willing to manufacture solar modules from quartz would be a good move by the government to encourage domestic manufacturing and also ensure energy security. We recommend levying BCD on solar wafers from 2027 to 2030 (in line with BCD on solar cells and modules) and the process to manufacture modules from quartz could be planned to begin by the same year. These upstream PV manufacturing stages are capital- and energy-intensive and the government may provide incentives to accelerate their scaling up. Support in the form of concessional power tariffs, capex subsidy and interst subvention may be critical for energy and capital intensive manufacturing stages such as polysilicon and ingot manufacturing. (Table 5)

Table 5 Recommended deployment targets for polysilicon and wafer manufacturing

Target year (cumulative capacity)	Quartz to polysilicon (tons/year)	Polysilicon to wafer (GW)
FY 2021–22	—	—
FY 2022–23	—	2
FY 2023–24	—	4
FY 2024–25	—	6
FY 2025–26	—	8
FY 2026–27	30,000	10
FY 2027–28	60,000	15
FY 2028–29	90,000	20
FY 2029–30	90,000	25

Source: CEEW-CEF analysis

Note: Polysilicon consumption is expected to decrease to 3–3.5 g/watt in the years to follow from the current values of 4 g/watt (Fraunhofer ISE 2020). A 30,000 tons/year facility may be capable of supplying to 8–10 GW of wafer manufacturing facility. The figure of 30,000 tons/year has been suggested keeping in mind the economies of scale to set up such facilities

The government should encourage public sector undertakings (PSUs) such as Bharat Heavy Electricals Limited (BHEL) and Bharat Electronics Limited (BEL) to take a lead in the solar manufacturing indigenisation effort. Any decision on this front should also involve stakeholders from other industries (such as mobile phones), which require polysilicon as a raw material.

(b) Horizontal integration

CEEW-CEF recommends the Ministry of Commerce and Industry (MoC&I) and Ministry of Electronics and Information Technology (MeitY) to implement this intervention

Currently, Indian manufacturers not just import cells but rely heavily on imports for other non-cell BOM components such as glass, ribbon, EVA sheet, and others, which constitute 30 to 35 per cent of the total module cost. Indian micro, small, and medium enterprises (MSMEs) are already producing these components, but they don't measure up on quality and are priced high as well. Currently anti-dumping duty (ADD) is levied on solar glass and EVA sheets in India. Despite the government's protective policy, domestic manufacturers capable of manufacturing these equipment have not ventured into the sector. A detailed analysis of each BOM component is required from the Indian MSMEs point of view and corrective measures should be urgently implemented, which can ensure the growth of the sector, create additional jobs, and increase competitiveness with China.

(c) Setting up manufacturing parks

CEEW-CEF recommends the Ministry of Commerce and Industry (MoC&I) and Ministry of Electronics and Information Technology (MeitY) to implement this intervention

A solar power plant does not have many moving parts. So manufacturing of key equipment like solar modules, inverters, and module mounting structure is easy and scalable. Imports is the source of many of the components and raw materials in equipment manufacture. To indigenise the manufacture of key equipment used in the solar power plant, we recommend that the government set up 10 GW solar manufacturing parks near the ports in the states of Maharashtra, West Bengal, Andhra Pradesh, and Tamil Nadu. Infrastructure in the form of assured and low-cost electricity, availability of water, concessional land, and air connectivity need to be ensured in these locations. Tax holiday and duty benefits may also be extended by the government. State govenrments must be intiate the process to skill the local workforce to work in such facilities. These parks can follow the model of SEZ but they need to be treated on par with DTA to avoid the levy of SGD or BCD on the manufactured products. Wind turbine and component manufacturers may also be given opportunities to set up their factories in these parks.

(d) Indigenisation of novel technologies

CEEW-CEF recommends the Ministry of New and Renewable Energy (MNRE) and Department of Science and Technology (DST) to implement this intervention

European countries have demonstrated success in the adoption of new technologies like organic solar modules from the lab to factory by commercialising innovative technologies that were initially developed by researchers in the lab. Prestigious technology institutions such as the Indian Institute of Technology (IIT) Bombay, IIT Delhi, IIT Madras, and the Indian Institute of Science (IISc) Bangalore should be roped in by the government, by providing financial grants, to develop and commercialise novel technologies, thus making India self-reliant.

Some interventions only yield short-term benefits: tread with caution

Support in the form of upfront capex support, interest subvention, assured off-take, and manufacturing-linked tenders have either been implemented before or may be in process of deployment. They have not been mentioned as probable solutions for the growth of domestic solar manufacturing due to the following reasons.

(a) **Regulations should support technological advancement**—Technological developments are taking place very rapidly across the world, and bankers and policymakers are not expected to be up-to-date on latest technologies. The government may end up supporting obsolete technologies due to lack of information. Besides, significant delays have been seen in the past for capex support and manufacturing-linked tenders, which can have negative impact on the sector. Due to delay, the window of opportunity to set up any facility also gets reduced significantly.

To avoid the trap of lack of expertise, bankers should hire experienced consultants who can provide expert assistance with latest technologies for evaluating solar manufacturing proposals. Many manufacturers have aired the view that banks are not comfortable lending to cell manufacturing facilities. This is purely due to lack of knowledge on the part of bankers. We recommend that the Indian Renewable Energy Development Agency (IREDA) be appointed as the lead agency for banks and policymakers to provide assistance in assessing the technical feasibility of solar manufacturing project proposals.

(b) **Availability and cost of capital is always a challenge**—Availability and cost of capital has remained a recurring problem for India's RE sector. However, RE project developers have managed to secure low-cost capital (from India and abroad) through financial engineering. If favourable policies are in place, RE manufacturers would also be able to raise low-cost capital to scale up their facilities. The government should go a step further in creating a vibrant solar manufacturing ecosystem where entrepreneurs can innovate and make course corrections based on market conditions and requirements. Developing an effective ecosystem to attract entrepreneurs and investment in solar manufacturing hinges on providing visibility of future deployment, policy stability and continuity, quality infrastructure, and a mature banking system.

(c) **Disbursement delays and decisions**—Disbursement of subsidies by the government often encounter delay. Factoring in the risk of such delay, manufacturers are burdened with increased cost, reducing the impact of fiscal support. Additional paperwork and bureaucratic hurdles add to the frustration of the manufacturers, giving rise to suspicions and possibility of foul play. To overcome these issues, the government should stipulate a timeline for each intervention.

In addition, overlapping mandates of different ministries like MNRE, Ministry of Commerce and Industries (MoC&I), Ministry of Finance (MoF), and Ministry of Electronics and Information Technology (MeitY) have often resulted in delays and sub-optimal outcomes. A steering committee may be formed to coalesce the efforts of various ministries for effective implementation of policies.

(d) **Mixing interventions can have the opposite effect**—The manufacturing-linked tender mixes deployment with manufacturing, which have two unrelated business processes. This complexity reduces market participation in the bids, leading to discovery of non-optimal prices and additional burden on the exchequer. Targeting low-cost deployment and low-cost manufacturing simultaneously will always result in a non-optimal outcome. Challenges for each sector are unique and the government should derive outcomes based on informed decision-making. Merely merging several issues in the expectation of an optimal outcome only ends up burdening the end consumers and discoms.

(e) **Expectation from others—**Expecting autonomous government institutions to continuously set up projects based on DCR is a misplaced strategy. While autonomous institutions such the Indian Railways, oil management companies, and NTPC Limited (formerly National Thermal Power Corporation) can set up some projects, the additional cost would be passed on to their consumers. Hence, while they may be encouraged to buy power from DCR projects (implemented following the recommended VGF scheme in this report), they cannot be burdened with the full responsibility to scale up the domestic manufacturing industry.

5. Conclusion

Indian PV module manufacturers lack competitiveness because of higher prices and lower capacity utilisation in comparison to their Chinese counterparts. The increased cost of Indian modules arises from higher raw material and overhead costs. At full capacity utilisation levels, BOM costs alone account for 56 per cent of the price difference between Indian and Chinese modules. India's manufacturing base also remains so low that economies of scale cannot be realised.

An integrated cell and module manufacturing generates around 2.6 full-time equivalent jobs per MW of output. Every 10 GW of additional cell and module manufacturing capacity could generate 26,000 jobs in the PV cell manufacturing sector (Kuldeep 2017). Additional jobs can be created in ingot and wafer manufacturing. Completely relying on local components and raw materials saves forex outflow. Domestic manufacturers also stand to benefit in the long term if they scale up and deploy superior technologies. They can then break into the international market as well by initially supplying modules to member countries of the International Solar Alliance (ISA).

However, to scale up domestic manufacturing in the country, a mix of well thought-out and long-term interventions are needed that can not only support the sector but also make India export competitive and a global leader in solar PV manufacturing technology.

For the domestic industry to gain momentum, the government needs to articulate policies with clear timelines for infusing confidence in the manufacturers and to attract investment. In order to boost the competitiveness of Indian manufacturing, measures that increase capacity utilisation and facilitate scaling up of horizontal and vertical integration can be effective in lowering selling prices. Recognising the solar manufacturing sector as strategic for country's growth and setting up an inter-ministerial body of senior officers may be the first step to achieve self-reliance as well become globally competitive.

The manufacturers would benefit by seeking support for clear mandates with specific timelines and keeping pace with the dynamically evolving nature of the sector. Manufacturing in a globalised world requires strategic planning and informed decision-making. The Indian government has already provided a long-term commitment to scaling up RE, which can be leveraged effectively to develop a thriving manufacturing base in India. Reduced reliance on imported products will make the sector self-sufficient, competitive, and resilient to supply chain disruptions. Apart from increasing India's energy security, the domestic solar manufacturing industry would also help in the creation of jobs and growth of the economy.

References

CBIC. 2017. "General Exemption No. 183." CBIC. 30 June. Accessed June 25, 2020. https://www.cbic.gov.in/resources//htdocs-cbec/customs/cst1718-020218/G.E.%20183.pdf.

Fraunhofer ISE. 2020. "Photovoltaics Report." Fraunhofer ISE. 17 June. Accessed June 24, 2020. https://www.ise.fraunhofer.de/content/dam/ise/de/documents/publications/studies/Photovoltaics-Report.pdf.

Hutchins, Mark. 2019. "The Weekend Read: Size Matters." October 5. Accessed June 25, 2020. https://www.pv-magazine.com/2019/10/05/the-weekend-read-size-matters/.

IRENA. 2019. "Future of Solar Photovoltaic." IRENA. November. Accessed June 2020. https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2019/Nov/IRENA_Future_of_Solar_PV_2019.pdf.

LONGi. 2020. "Seven Leading Photovoltaic Industry Companies Jointly Establish the M10 Silicon Wafer Standard Size." June 24. Accessed June 26, 2020. https://en.longi-solar.com/home/events/press_detail/id/245.html.

MNRE. 2020. "MNRE Extends Effective Date for Implementation of Approved Lists of Models and Manufacturers of Solar PV Modules and Solar PV Cells by Six Months to 30.09-2020." MNRE. April. Accessed June 2020. https://mnre.gov.in/img/documents/uploads/file_f-1587720242828.pdf.

MNRE. 2020a. "CPSU Scheme." MNRE. 13 April. Accessed June 25, 2020. https://mnre.gov.in/img/documents/uploads/file_f-1586766335590.pdf.

MNRE. 2020b. "MNRE." MNRE. 12 February. Accessed June 2020. https://mnre.gov.in/img/documents/uploads/file_f-1582100822990.pdf.

MNRE. 2020c. "Solar PV Manufacturing." 25 June. Accessed June 2020. https://mnre.gov.in/solar/manufacturers-and-quality-control.

MoF. 2020. "Notification No. 02/2020-Customs (SG)." CBIC. 29 July. Accessed July 30, 2020. https://www.cbic.gov.in/resources//htdocs-cbec/customs/cs-act/notifications/notfns-2020/cs-sg2020/cssg02-2020.pdf;jsessionid=02979DDFE8BADE1A651ED5BE408A27E4.

Kuldeep, Neeraj, Kanika Chawla, Arunabha Ghosh, Anjali Jaiswal, Nehmat Kaur, Sameer Kwatra, Karan Chouksey. 2017. *Greening India's Workforce Gearing up for Expansion of Solar and Wind Power in India*. June. Accessed June 2020.

WTO. 2016. "India — Certain Measures Relating to Solar Cells and Solar Modules." WTO. October. Accessed June 25, 2020. https://www.wto.org/english/tratop_e/dispu_e/cases_e/ds456_e.htm.

WTO. 2020. "Information Technology Agreement News Archive." Accessed June 25, 2020. https://www.wto.org/english/news_e/archive_e/ita_arc_e.htm.

Xiao, Carrie. 2020. "'Never Stop Innovating': Jinko CTO Talks R&D and Solar's Future Direction." 24 March. https://www.pv-tech.org/editors-blog/never-stop-innovating-jinko-cto-talks-rd-and-solars-future-direction.

Annexures

Annexure I

Impact matrix of the key interventions in the solar manufacturing sector

S. No.	Intervention	Short-term/ long-term	Fiscal implication (for government)	Ease of implementation	Impact on the sector
1	Tariff barrier	Long-term	Inflow	Easy	Survival
2	Production subsidy	Short-term	Outflow	Difficult	Survival
3	VGF	Long-term	Outflow	Moderate	Growth
4	Manufacturing equipment (duty exemption)	Medium-term	Outflow	Easy	Moderate
5	Commitment from the government	Long-term	Nil	Easy	Growth
6	Better recording of data	Short-term	Nil	Easy	Growth
7	Backward integration	Long-term	Outflow	Difficult	Growth
8	Horizontal integration	Medium-term	Outflow	Difficult	Moderate
9	Indigenisation of novel technologies	Medium-term	Outflow	Easy	Growth
10	Manufacturing parks	Medium-term	Outflow	Easy	Moderate

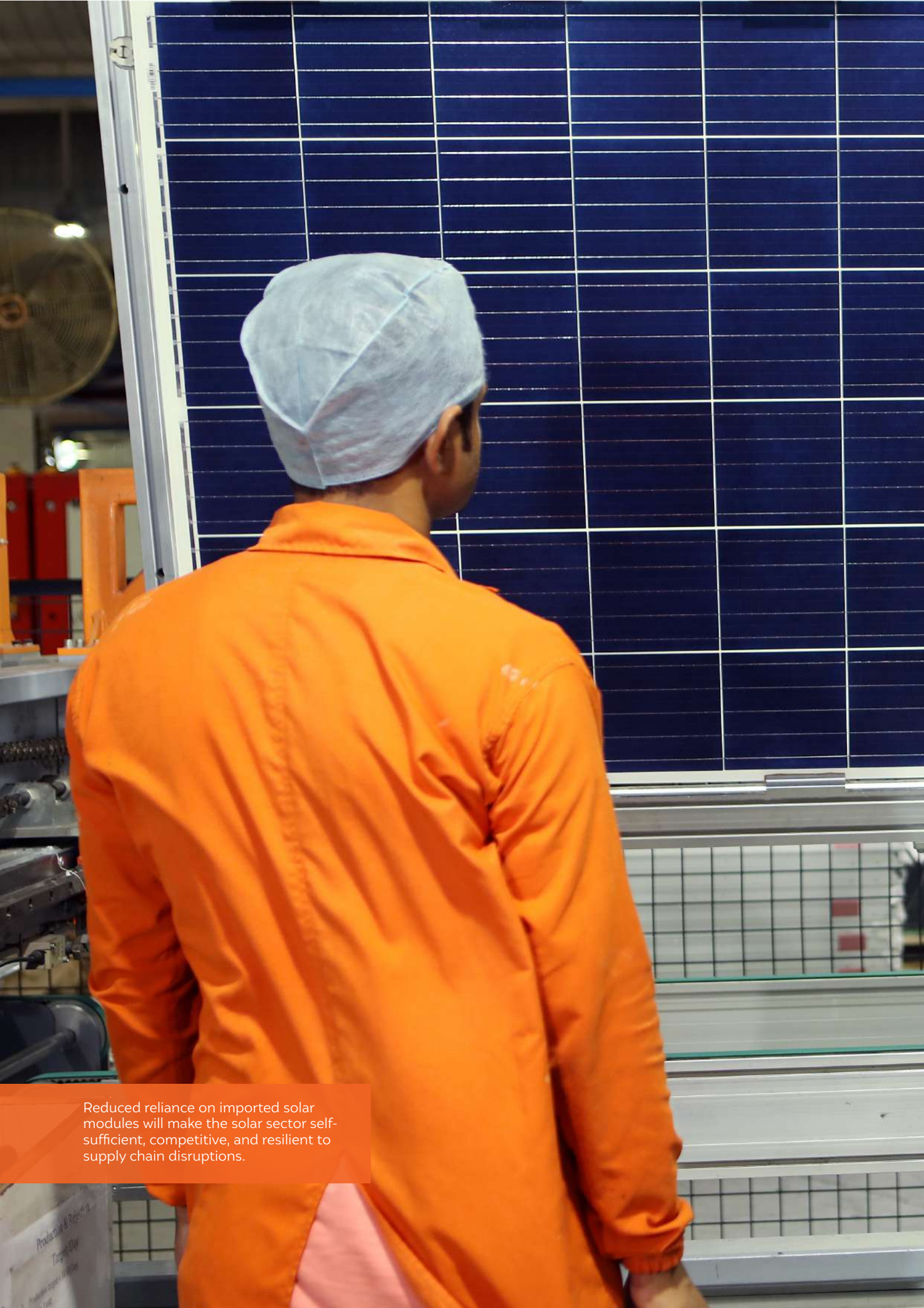
Source: CEEW-CEF analysis

Annexure II

Key respondents to primary survey

Solar Module manufacturers approached	Name of respondents with capacity
Adani Solar	Adani Solar (1,200 MW)
Emmvee Solar	
Goldi Solar Private Limited	Goldi Solar (500 MW)
IB Solar	
Jakson Limited	Jakson Limited (90 MW)
Navitas Green Solutions Private Limited	Navitas Green Solutions Private Limited (200 MW)
Renewsys India Private Limited	Renewsys India Private Limited (700 MW)
Tata Power Solar	
Vikram Solar Limited	Vikram Solar Limited (1,000 MW)
Waaree Energies Limited	Waaree Energies Limited (2,000 MW)

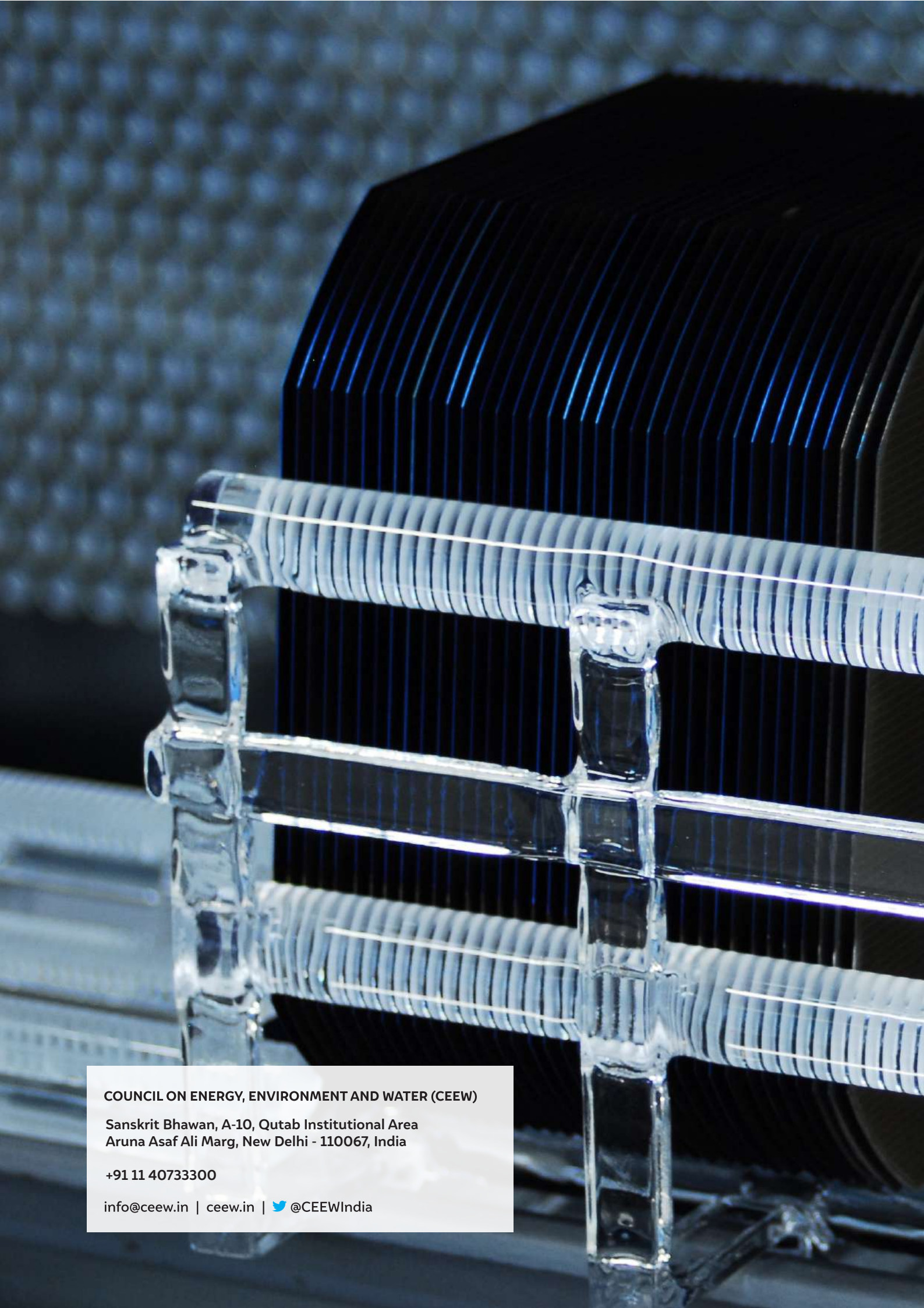
Source: CEEW-CEF compilation



Reduced reliance on imported solar modules will make the solar sector self-sufficient, competitive, and resilient to supply chain disruptions.



Image: Vikram Solar ltd



COUNCIL ON ENERGY, ENVIRONMENT AND WATER (CEEW)

Sanskrit Bhawan, A-10, Qutab Institutional Area
Aruna Asaf Ali Marg, New Delhi - 110067, India

+91 11 40733300

info@ceew.in | ceew.in | @CEEWIndia

2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)
Amity University, Noida, India. June 4-5, 2020

Software Defect Categorization based on Maintenance Effort and Change Impact using Multinomial Naïve Bayes Algorithm

Ruchika Malhotra
Department of Computer Science and Engineering,
Delhi Technological University
Shahbad Daulatpur, Bawana Road, Delhi, India
ruchikamalhotra2004@yahoo.com

Madhukar Cherukuri
Department of Computer Science and Engineering,
Delhi Technological University
Shahbad Daulatpur, Bawana Road, Delhi, India
srkr.madhu@gmail.com

***Abstract*— Cost-to-fix software defects increase exponentially as the software moves forward in its life cycle. The fixing of defects found during the maintenance phase is very costly. Software Defect Categorization (SDC) models categorize the software defects into different levels (high, moderate and low) based on various attributes. Such models can aid the software maintenance team to prioritize the defects and plan the resources effectively in fixing them. This study develops SDC models based on three software bug attributes – i) maintenance effort needed to fix a defect ii) change impact on the software for fixing a defect and iii) the combined approach of both - maintenance effort and change impact. In this study, we extracted the important features from the defect reports using text mining and the classification models are developed for each of the attributes using Multinomial Naïve Bayes (NBM) algorithm. The capability of the models is calculated using Area Under the ROC (Receiver Operating Characteristic) curve (AOC). The results of the study show that i) SDC models based on maintenance effort and change impact are capable to categorize the software defects ii) the performance of the SDC models based on the combined approach is better than SDC models based on maintenance effort and change impact.**

Keywords—Software Defect Categorization, Software Maintenance, Change Impact, Maintenance Effort, Multinomial Naïve Bayes, Machine Learning

I. INTRODUCTION

Software maintenance is an important phase of the software development life cycle. It involves fixing defects that are encountered in the production, implementing the new requirements and change requests of the customer. The changes to the software in production are real-time and hence should be implemented and tested in a time sensitive manner with the limited available resources.

Software Defect Categorization (SDC) is found to be useful in managing the resources effectively to fix defects during the maintenance phase. It involves the assignment of a level (high, medium, low) to a software defect based on the defect attributes such as their criticality, priority, and others [2]-[4].

The existing literature on the defect categorization has shown that the assignment of the defect level based on maintenance effort and change impact have good prediction capability compared to the categorization based on the other defect attributes. In this study, we have explored the combined approach of maintenance effort and change impact to build the models that categorize defects and their performance is evaluated and compared with models based on maintenance effort and change impact.

The maintenance effort needed to fix a defect is estimated by the number of lines of code (LOC) changed (added or deleted or altered) to remove a software defect. The change impact of a software defect is the number of classes that are changed to fix the corresponding defect. Software defects categorized as high level based on the maintenance effort requires more resources for removing the defect. Those defects categorized as high level based on the change impact requires more resources for regression testing. Those categorized as high based on both attributes require more resources for the removal of defects and regression testing.

Our study extracts the keywords by mining the textual descriptions of the error reports, builds SDC Models using Multinomial Naïve Bayes (NBM) learning algorithm, and the categorization capability of these models is assessed using AOC measure. The experiments were conducted on five modules of the Android operating system (Bluetooth, Browser, Calendar, Camera and MMS).

The paper is organized into the following sections: Section II describes the related research. The background of our work is discussed in Section III. Section IV gives a brief description of the classification algorithm and performance measure employed in our work. The experimental outputs of the work are presented in Section V. Section VI analyses and discuss the results of the work. The possible threats to the validity of our work are discussed in Section VII. Finally, Section VIII concludes the study and presents future work.

II. RELATED WORK

Software Researchers have conducted many studies to estimate the severity of defects based on their defect reports. Lamkanfi et al. used the textual description given in a defect report to estimate the criticality of defects [1]. Han et al. applied text mining and deep learning on the defect description to estimate the criticality of the corresponding defect [5]. Sharma et al. classified the bugs into two levels namely critical and non-critical [6]. Their study showed the performance of bug categorization based on the text mining of the bug reports. Chaturvedi et al. conducted a study to evaluate the effectiveness of six classification algorithms to

978-1-7281-7016-9/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 21,2020 at 05:09:21 UTC from IEEE Xplore. Restrictions apply.

estimate the criticality of a defect based on its textual description [7]. Yang et al. evaluated three parameter tuning methods namely information gain, chi square and correlation coefficient with multinomial naïve bayes classifier to predict defect criticality based on the bug reports [8].

Researchers have conducted studies to analyze the change impact by static code inspection [9][10], dynamic code inspection [11] or the combination of both[12]. Some studies have used testing mining techniques on the version histories to analyze the change impact on the software [13][14].

Most of the researchers in their studies used LOC to predict the software maintenance effort estimation [15]-[18]. Balogh et al. [19] estimated the maintenance effort based on software metrics such as Chidamber & Kemerer metrics suite [20], programmer's skills, logical LOC, coding effort and others. Kumar et al. [21] calculated the modified LOC between two versions and categorized their target variable into low, medium and high levels.

III. Research Background

The study analyzes the defect data of five modules of Android software [22], that is available on GITHUB repository [23]. The changelogs between two versions in the repository have a description of the changes, that are used in finding the defects. The five Android modules and their versions between which change logs extracted were: Bluetooth (4.1-4.4), Browser (2.3-4.0), Calendar (4.1-4.4), Camera (2.3-4.0) and MMS (2.3-4.0).

The study used software called Defect Collection and Reporting System (DCRS), that aids in extracting the bug reports from the changelogs between the two predetermined versions of Git repository [24]. Each bug report extracted has the following data: a) Unique defect identifier; b) Source code file name, which has been affected by the defect correctness; c) the description of the defect; d) LOC added and LOC removed to fix a defect and e) the total of LOC added or removed for fixing the defect.

We applied text mining techniques to find keywords from the defect descriptions found in the bug reports generated by DCRS. We performed preprocessing, feature selection using infogain measure and vector space model defined by Ruchika et al [28][29] to obtain Top10, Top25, Top50 and Top100 ranked keywords. These keywords are used in building SDC models as independent features. We categorize the defects found from the bug reports based on maintenance effort and change impact and assign levels (low, medium and high) for every defect identified in the android application packages under this study.

We build Software Defect Categorization (SDC) models using NBM classification algorithm with ten-fold cross validation [25]. The dataset extracted from the bug reports is equally divided into ten parts, where nine parts are used as a training set while building the model and the remaining tenth part for validating it. This process is repetitively performed ten times until all the divided ten subsets have been used for validating the model once. Three SDC models (low level, medium level, and high level) using NBM classification algorithm are developed for each dataset under study.

IV. Research Methodology

We used Multinomial Naïve Bayes classification algorithm to build SDC Model and its performance is measured using Area under Receiver Operating Characteristics Curve(AUC). Weka tool with its default parameter configuration is used to build the SDC Models [26]. In this section, we have given a brief description of the NBM classification technique and AUC performance measure used in the study.

A. Naïve Bayes Multinomial

A naive bayes classifier assumes that each of the variables it uses are conditionally independent of one another given some class. More formally, to calculate the probability of observing features f_1 through f_n, given some class c, under the naive bayes assumption the following holds:

$$p(f_1,..,f_n/c) = \prod_{i=1}^{n} p(f_i/c)$$

This means that when we want to use a Naive Bayes classifier to classify a new example, the posterior probability is much simpler to work with:

$$p(c/f_1,..,f_n) \ \alpha \ p(c)p(f_1/c)...p(f_n/c)$$

These assumptions of independence are rarely true, which may explain why some have referred to the model as the "Idiot Bayes" model, but in practice, Naive Bayes models have performed surprisingly well, even on complex tasks where it is clear that the strong independence assumptions are false.

The term Multinomial Naive Bayes means that each $p(f_i/c)$ is a multinomial distribution and holds good for ordinal values. In summary, NBM model is a special case of Naive Bayes model where a multinomial distribution is used for each of its features.

B. Performance Measures

We used AUC for measuring the performance of the SDC models as it proved to be a stable method [27][29]. We now define sensitivity and specificity for these models. The sensitivity of a "low" level model is the ratio of correctly classified "low" level data points as to the total number of actual "low" level data points. The specificity of a "low" level model is the ratio of correctly classified "not low" level data points to the total number of actual "not low" level data points. The sensitivity and specificity of "medium" level and "high" level models can be defined similarly. Receiver Operating Characteristic (ROC) curve is a plot between sensitivity on the y-coordinate and 1-specificity on the x-coordinate. The area under ROC Curve (AUC) value of 1 indicates that the model prediction is 100% accurate and hence the models with higher values of AUCs are desired.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 21,2020 at 05:09:21 UTC from IEEE Xplore. Restrictions apply.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 21,2020 at 05:09:21 UTC from IEEE Xplore. Restrictions apply.

TABLE I : AUC values of SDC Models based on Maintenance Effort

Defect Level	Bluetooth				Browser			Calendar				Camera				MMS			
	Top-10	Top-25	Top-50	Top-100	Top-10	Top-25	Top-50	Top-5	Top-10	Top-25	Top-50	Top-10	Top-25	Top-50	Top-100	Top-10	Top-25	Top-50	Top-100
Low	**0.753**	**0.799**	**0.913**	**0.874**	0.585	0.597	0.653	0.569	0.681	**0.818**	**0.837**	0.61	**0.702**	**0.759**	**0.826**	0.421	**0.72**	**0.77**	**0.834**
Medium	**0.75**	**0.833**	**0.868**	**0.891**	0.574	0.604	0.669	0.482	0.563	**0.741**	**0.788**	0.626	0.694	**0.724**	**0.815**	0.448	0.6	0.641	**0.726**
High	0.687	**0.833**	**0.825**	**0.881**	0.573	0.54	0.604	0.621	0.695	**0.871**	**0.9**	**0.713**	**0.795**	**0.805**	**0.853**	0.494	0.643	0.673	**0.774**

TABLE II : AUC values of SDC Models based on Change Impact

Defect Level	Bluetooth				Browser			Calendar				Camera				MMS				
	Top-10	Top-25	Top-50	Top-100	Top-10	Top-25	Top-50	Top-100	Top-10	Top-25	Top-50	Top-100	Top-10	Top-25	Top-50	Top-100	Top-10	Top-25	Top-50	Top-100
Low	0.577	0.68	0.609	0.513	0.567	0.554	0.566	0.643	0.588	**0.724**	**0.774**	**0.801**	0.629	0.672	**0.716**	**0.759**	0.535	0.541	0.644	0.632
Medium	0.489	0.592	0.607	0.581	0.511	0.476	0.502	0.517	0.487	0.591	0.557	0.562	0.514	0.592	0.586	0.627	0.615	0.625	0.642	0.546
High	0.418	0.518	0.67	0.665	0.557	0.536	0.552	0.662	0.66	**0.812**	**0.783**	**0.812**	0.652	**0.743**	**0.756**	**0.765**	0.367	0.552	0.593	0.678

TABLE III : AUC values of SDC Models based on Maintenance Effort and Change Impact

Defect Level	Bluetooth				Browser				Calendar				Camera				MMS			
	Top-10	Top-25	Top-50	Top-100	Top-10	Top-25	Top-50	Top-100	Top-10	Top-25	Top-50	Top-100	Top-10	Top-25	Top-50	Top-100	Top-10	Top-25	Top-50	Top-100
Low	**0.739**	**0.807**	**0.908**	**0.9**	0.575	0.586	0.625	**0.705**	0.643	**0.81**	**0.801**	**0.848**	0.638	0.67	**0.755**	**0.794**	0.67	**0.735**	**0.819**	**0.871**
Medium	0.664	**0.796**	**0.714**	**0.731**	0.578	0.57	0.624	0.658	0.571	**0.713**	**0.747**	**0.815**	0.523	0.635	**0.707**	**0.766**	0.676	**0.782**	**0.836**	**0.861**
High	0.612	**0.807**	0.698	0.683	0.545	0.551	0.574	0.688	0.654	**0.819**	**0.882**	**0.889**	0.672	**0.784**	**0.799**	**0.846**	0.655	**0.753**	**0.768**	**0.87**

V. EXPERIMENTAL RESULTS

A. SDC Model using Maintenance Effort

The AUC values of the SDC models developed in accordance with Maintenance effort using the NBM classification technique on the five datasets used in the study are shown in the Table I. The developed SDC models identify three specific levels allocated to software defects i.e. "low", "medium" or "high". We develop four models at each level using a different number of predictor variables (Top 10, Top 25, Top 50 and Top 100). All the AUC values which are greater than 0.7 are marked in bold. It is evident from the table I that the ranges of AUC values of the "low level" SDC models of Bluetooth, Browser, Calendar, Camera and MMS datasets ranged from 0.753-0.913, 0.585-0.653, 0.569-0.837, 0.610-0.826 and 0.421-0.834 respectively. The ranges of AUC values obtained by "medium level" SDC models were 0.750-0.891 (Bluetooth), 0.574-0.669 (Browser), 0.482-0.788 (Calendar), 0.626-0.815 (Camera) and 0.448-0.726 (MMS). Similarly, the ranges of AUC value for "high level" SDC models 0.687-0.881, 0.540-0.604, 0.621-0.900, 0.713- 0.853, and 0.494-0.774 for the Bluetooth, Browser, Calendar, Camera and MMS datasets respectively. These ranges of AUC values indicate that the categorization of defects into different levels according to the SDC models developed using the maintenance effort required to correct them is acceptable and accurate.

An analysis of the AUC values of "low level", "medium level" and "high level" SDC models indicate that the "high level" category models obtained higher AUC and accuracy values on three of the datasets (Browser, Calendar, and Camera). As mentioned earlier, it is important to detect "high level" defects according to the maintenance effort so that maintenance team managers can allocate appropriate resources for correcting these defects. This will help in reducing the cost to fix defects and so effective maintenance of the software with optimized resource usage.

The AUC values for "low level" SDC models ranged from 0.753-0.913, 0.585-0.653, 0.569-0.837, 0.610-0.826 and 0.421-0.834, while the AUC values of "medium level" SDC models were found to be in the range of 0.750-0.891, 0.574-0.669, 0.482-0.788, 0.626-0.815 and 0.448-0.726 in most of the cases in Bluetooth, Browser, Calendar, Camera and MMS datasets respectively. The AUC values for "high level" SDC models were found to be in the range of 0.687-0.881, 0.540-0.604, 0.621-0.900, 0.713- 0.853, and 0.494-0.774 respectively in Bluetooth, Browser, Calendar, Camera and MMS datasets respectively. These AUC values indicate that the categorization of defects into different levels according to the corresponding maintenance effort required to correct them is acceptable and accurate.

B. SDC Model using Change Impact

The AUC values of SDC models developed in accordance with Change Impact using the NBM classification technique on the android data sets used in the study are shown in Table II. We marked all AUC values with a value greater than 0.7 as bold in Table II. The AUC values for "low level" SDC models ranged from 0.513-0.680, 0.554-0.643, 0.588-0.801, 0.629-0.759 and 0.535-0.644, while the AUC values of "medium level" SDC models were found to be in the range of 0.489-0.607, 0.476-0.517, 0.487-0.591, 0.514-0.627 and 0.546-0.64 in most of the cases in Bluetooth, Browser, Calendar, Camera and MMS datasets

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 21,2020 at 05:09:21 UTC from IEEE Xplore. Restrictions apply.

respectively. The AUC values for "high level" SDC models were found to be in the range of 0.418-0.665, 0.536-0.662, 0.660-0.812, 0.652-0.765 and 0.367-0.678 respectively in Bluetooth, Browser, Calendar, Camera and MMS datasets.

As seen in the table II, most of the AUC values are accurate and acceptable, but, few SDC models obtained AUC values in the range 0.3-0.5. These poor values were obtained as the SDC models were developed from highly imbalanced training data.

C. SDC Model according to the combined effect of its Maintenance Effort and Change Impact

SDC models were developed using the NBM classification technique in accordance with the product of a bug's required maintenance effort and its change impact values. Table III states the AUC results of the developed SDC models. All AUC values greater than 0.7 are marked in bold.

As shown in the table III, the SDC models built at "high" level category obtained AUC values in the range 0.612-0.807, 0.545-0.688, 0.654-0.889, 0.672-0.846 and 0.655-0.870 in most of the cases for Bluetooth, Browser, Calendar, Camera and MMS datasets respectively. The SDC models at "low level" obtained the ranges of AUC values of from 0.739-0.908, 0.575-0.705, 0.643-0.848, 0.638-0.794 and 0.670-0.871, while those at "medium level" obtained the ranges of AUC values of 0.664-0.796, 0.570-0.658, 0.571-0.815, 0.523-0.766 and 0.676-0.861 respectively for Bluetooth, Browser, Calendar, Camera and MMS datasets. Most of the obtained AUC values by the SDC models depicted in Table III were acceptable.

VI. Discussion of Results

The results of our work show that the AUC values of the SDC models built using the combination of maintenance effort and change impact was comparable to SDC models built in accordance with the maintenance effort. As the combined approach considers both the maintenance effort and the change impact values, they found to be more useful. Maintenance effort aids in planning the developer's effort based on the level of defects predicted, while the change impact aid in planning the tester's effort based on the prediction of the number of the classes that are changed to fix the defects. Testers can perform regression testing on only the modules that are impacted, which saves cost and time.

The results of the study also show that the AUC values of the SDC models built using the combined approach were better than SDC models built in accordance with change impact values. The primary reason for the poor performance of SDC models based on change impact values was the imbalanced data of the training dataset.

It can also be observed from the results of the study that all the performance of the developed SDC models for Top-10 features was poor as compared to Top-25, Top-50, and Top-100 features. In most of the instances, the AUC of SDC models developed using Top-10 keywords were lesser than Top-25, Top-50 or Top-100 words. The primary reason for such poor performance was the inability of Top-10 words to encapsulate the required information for predicting the level of a defect based on its bug report. On the other hand, both the SDC models developed using Top-50 and Top-100 keywords were found appropriate with effective AUC values. The reason for the comparable performance of Top-50 and Top-100 models could be redundancy in the predictors. There is a possibility that Top-100 words may be representing redundant information in certain cases, which was effectively encapsulated in just Top-50 words. In such cases, developing models with Top-50 words is more.

VII. Threats to Validity

The results of the our developed SDC models assume that there exists a relationship between the defect description of the bug reports and their corresponding maintenance effort and change impact values. Studies in the existing literature have also held such relationships [31][32]. Some studies showed the relationship between descriptions of a defect report and other defect factors such as priority or severity, which further strengthens our assumption of such relationships [1]-[8].

As each dataset has defect reports of three possible categories ("low", "medium," & "high") and the SDC models built are binary, the training data is imbalanced. Various studies showed that AUC is a good measure for assessing the performance of the models trained with imbalanced data [27]-[30]. By using AUC, we mitigated the threat of imbalanced data.

The predictions of our models have been empirically validated using defect reports from the five datasets of Android modules and hence the results may not be generalized. Various studies show that the results are widely applicable in similar domain open-source software datasets. Future research may be carried out on the defect reports from other application domains.

VIII. Conclusion and Future Work

This study developed SDC models based on three defect attributes, i.e. its maintenance effort, its change impact and the combined effect of both maintenance effort and change impact. These three models built using each of these defect attributes, which are capable to estimate whether a software defect belongs to "low category" or "not low category", "medium category" or "not medium category", "high category" or "not high" category. The categorization is done by text mining the relevant keywords from defect reports. Four models are developed using the NBM classification technique for each scenario using Top 10, Top 25, Top 50 and Top100 relevant keywords for five Android Operating System Modules. Experimental results of our work are validated using the AUC values of the SDC models. The results of our study are as follows:

- The results of AUC values of "high" category SDC models on all the datasets used in the study for Top 25, Top 50 and Top 100 words as predictors are summarized as follows: The most of AUC values of SDC models built in accordance of maintenance effort are in the range 0.643-0.881 for all the datasets of the study. The most of AUC values of SDC models built based on change impact values were in the range 0.652-0.812. The most of AUC values of SDC models built based on the product of maintenance effort and change impact are in the range of 0.654-0.889. These AUC values indicate acceptable predictability of the SDC models built in our study. Similar values were shown in table II and

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 21,2020 at 05:09:21 UTC from IEEE Xplore. Restrictions apply.

table III for “medium” category and “low” category SDC models respectively.

- The performance of the SDC models according to the combination of maintenance effort and change impact are found to be better than the SDC models based on change impact and comparable to the SDC models according to maintenance effort.

Software practitioners can use the developed SDC models to estimate the developer and tester efforts and costs and ensure the optimal usage of limited resources. Software Project managers can use the combined approach SDC models to predict “high” category defects as they should be given a larger percentage of maintenance resources and adequate testing resources. The defects categorized as “high” in accordance with maintenance effort and “low” in accordance with change impact require more developer effort and less tester effort. Whereas, the defects categorized as “low” in accordance with maintenance effort and “high” in accordance with change impact require less developer effort and more tester effort.

Researchers may develop and validate the proposed models on other datasets of different domains and platforms and may validate the generalization of results. They may also use other classification algorithms for building the SDC models.

ACKNOWLEDGMENT

We would like to thank Dr. Megha Khanna Ummat for her inputs on the defect categorization framework and Delhi Technological University for its continuous support and encouragement.

REFERENCES

[1] Lamkanfi, A., Demeyer, S., Giger, E., & Goethals, B. Predicting the severity of a reported bug. in 7th IEEE Working Conference on Mining Software Repositories, (Cape Town, South Africa, 2010) IEEE, 1-10.

[2] Lamkanfi, A., Demeyer, S., Soetens, Q. D., & Verdonck, T. Comparing mining algorithms for predicting the severity of a reported bug. in 15th European Conference on Software Maintenance and Reengineering, (Oldenburg, Germany, 2011) IEEE, 249-258.

[3] Menzies, T., & Marcus, A. Automated severity assessment of software defect reports. in International Conference on Software Maintenance, (Beijing, China, 2008) IEEE, 346-355.

[4] Tian, Y., Lo, D., Xia, X., & Sun, C. (2015). Automated prediction of bug report priority using multi-factor analysis. Empirical Software Engineering, 20(5), 1354-1383.

[5] Han, Z., Li, X., Xing, Z., Liu, H., & Feng, Z. Learning to Predict Severity of Software Vulnerability Using Only Vulnerability Description. in International Conference on Software Maintenance and Evolution, (Shanghai, China, 2017) IEEE, 125-136.

[6] Sharma, G., Sharma, S., & Gujral, S. (2015). A Novel Way of Assessing Software Bug Severity Using Dictionary of Critical Terms. Procedia Computer Science, 70, 632-639.

[7] Chaturvedi, K. K., & Singh, V. B. Determining bug severity using machine learning techniques. In International Conference on Software Engineering, (Indore, India, 2012) IEEE, 1-6.

[8] Yang, C. Z., Hou, C. C., Kao, W. C., & Chen, X. An empirical study on improving severity prediction of defect reports using feature selection. in 19th Asia-Pacific Software Engineering Conference, (Hong Kong, China, 2012). IEEE, 240-249.

[9] Acharya, M., & Robinson, B. Practical change impact analysis based on static program slicing for industrial software systems. in International Conference on Software Engineering (Honolulu, Hawaii, 2011). ACM, 746-755.

[10] Jashki, M. A., Zafarani, R., & Bagheri, E. Towards a more efficient static software change impact analysis method. in ACM SIGPLAN-SIGSOFT workshop on Program analysis for software tools and engineering, (Atlanta, USA, 2008) ACM, 84-90.

[11] Cai, H., & Santelices, R. (2015). A comprehensive study of the predictive accuracy of dynamic change-impact analysis. Journal of Systems and Software, 103, 248-265.

[12] Maia, M. C. O., Bittencourt, R. A., de Figueiredo, J. C. A., & Guerrero, D. D. S. The hybrid technique for object-oriented software change impact analysis. In European Conference on Software Maintenance and Reengineering, (Madrid, Spain, 2010). IEEE, 252-255.

[13] Zimmermann, T., Zeller, A., Weissgerber, P., & Diehl, S. (2005). Mining version histories to guide software changes. IEEE Transactions on Software Engineering, 31(6), 429-445.

[14] Abdeen, H., Bali, K., Sahraoui, H., & Dufour, B. (2015). Learning dependency-based change impact predictors using independent change histories. Information and Software Technology, 67, 220-235.

[15] Jin, C., & Liu, J. A. Applications of support vector machine and unsupervised learning for predicting maintainability using object-oriented metrics. in International Conference on Multimedia and Information Technology, (Kaifeng, China, 2010), IEEE, 24-27.

[16] Van Koten, C. & Gray, A.R. (2006). An application of Bayesian network for predicting object-oriented software maintainability. Information and Software Technology, 48(1), 59-67.

[17] Malhotra, R. and Lata, K. An Exploratory Study for Predicting Maintenance Effort using Hybridized Techniques. in Innovations in Software Engineering Conference, (Jaipur, India, 2017), ACM, 26-33.

[18] Zhou, Y. & Leung, H. (2007). Predicting object-oriented software maintainability using multivariate adaptive regression splines. Journal of Systems and Software, 80(8), 1349-1361

[19] Balogh, G., Végh, Á.Z., & Beszédes, Á. Prediction of software development modification effort enhanced by a genetic algorithm. in Symposium of Search Based Software Engineering, (Trento, Italy, 2012). 1-6.

[20] Chidamber, S.R., & Kemerer, C.F. (1994). A metrics suite for object-oriented design. IEEE Transactions on software engineering, 20(6), 476-493.

[21] Kumar, L., & Sureka, A. Using structured text source code metrics and artificial neural networks to predict change proneness at code tab and program organization level. in Innovations in Software Engineering Conference, (Jaipur, India, 2017) ACM, 172-180.

[22] Source.Android.com, ‘Android Open Source Project (AOSP)’, 2007. [Online]. Available : https://source.android.com/. [Accessed : 26-Aug-2019]

[23] Github.com, ‘Software development version control using git’, 2008. [Online]. Available : https://github.com/github. /. [Accessed : 26-Aug-2019]

[24] Malhotra, R., Pritam, N., Nagpal, K., & Upmanyu, P. Defect collection and reporting system for git based open source software. in International Conference on Data Mining and Intelligent Computing (Delhi, India, 2014). IEEE, 1-7.

[25] Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. Journal of the royal statistical society. Series B (Methodological), 111-147.

[26] Weka 3, ‘The Workbench for Machine Learning’, 1994. [Online]. Available : http://www.cs.waikato.ac.nz/ml/weka/. [Accessed : 26-Aug-2019]

[27] He, H., & Garcia, E. A. (2009). Learning from imbalanced data. IEEE Transactions on knowledge and data engineering, 21(9), 1263-1284

[28] Malhotra, R. (2016). Empirical research in software engineering: concepts, analysis, and applications. CRC Press.

[29] Malhotra, R., & Khanna, M. (2017). An empirical study for software change prediction using imbalanced data. Empirical Software Engineering, 22(6), 2806-2851.

[30] Fawcett, T. (2006). An introduction to ROC analysis. Pattern recognition letters, 27(8), 861-874.

[31] Jindal, R., Malhotra, R., & Jain, A. Mining defect reports for predicting software maintenance effort. in International Conference on Advances in Computing, Communications and Informatics, (Kochi, India, 2015a) IEEE, 270-276.

[32] Jindal, R., Malhotra, R., & Jain, A. Predicting Software Maintenance effort using neural networks. in 4th International Conference on Reliability, Infocom Technologies and Optimization, (Noida, India, 2015b) IEEE, 1-6.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 21,2020 at 05:09:21 UTC from IEEE Xplore. Restrictions apply.

State-of-the-arts Violence Detection using ConvNets

Aayush Jain and Dinesh Kumar Vishwakarma

***Abstract*—Over the last few years, there has been a surge in developing automated systems that can monitor human activities in the public environment and recognize abnormalities like violent and suspicious events. Violence detection aims at identifying whether a violent action has occurred and has evolved as a popular topic in the field of image processing and computer vision. Effective techniques for accurate and intelligent video analysis are highly demanded. There are various methodologies which can detect such activities. This paper aims at reviewing various CNN based state-of-the-art techniques for violence detection. In this paper, we discuss variants of CNN, their application and datasets which play an indispensable part in the detection process.**

***Index Terms*—violence detection, CNN, image processing, computer vision**

I. Introduction

OVER the last decade, highly congested public places such as shopping centers, educational institutions, airports, hospitals, banks, markets, streets, etc. are being monitored by surveillance systems having CCTV camera and other equipment to monitor activities of people [1] for ensuring public safety and better crowd management due to rapid increase in global population. But the effectiveness of such systems is questionable since they require adequate manpower and human attention skills are limited [2]. Some assessments tell us that 99% of the generated footage is never watched [3] and hence, detection of suspicious activity is a very difficult task [4], [5]. True effectiveness is reduced to only 1% and hence the automated analysis of crowded scenes has received quite some attention from the research community. Researchers are now using computer vision techniques to recognize human activity [6]–[8] and monitor the crowd automatically. Different methodologies are already developed to filter out abnormalities in the crowd and helps to detect suspicious activity while monitoring cctv videos. Detecting violence by monitoring footage captured by CCTV and other surveillance videos [9], [10] is also a kind of activity detection.

Various algorithms, techniques and methods to detect violent behavior and actions and other harmful patterns in videos are surveyed by [11], [12]. In such methods, different approaches work with several parameters or features of videos such as optical flow [13], [14], acceleration, appearance, human pose, time, etc. Various researches talk about different techniques explored based on machine learning, support vector machines, deep learning to enhance the performance and better evaluation of the detection procedure by increased accuracy. This area of research toggles between violence detection and abnormal behavior detection, the activities which contain actions like fighting, beating, pushing, etc. are surveyed under violence detection. The primary objective of this research is to explore methods of violence detection using convolutional neural nets (CNN or ConvNets) and present a review of the same.

In this paper, we exhaustively explored and discuss CNN in section II and its different variants in section III, and in section IV we present review of previous work using a different or modified version of CNN used to detect violence in videos, their advantages, and limitations followed by exploration of real-world datasets available along with their characteristics and evaluation metrics used with such datasets, if any, in section V. Finally, we draw conclusion in section VI.

II. Convolutional Neural Network

An artificial neural network is made up of several simple, connected units called neurons, producing an arrangement of real-valued activations. Activation of neurons takes place through weighted links from the previous layer's active neurons. Neurons influence learning by forwarding the information. Learning is about obtaining weights that make neural networks exhibit the required behavior. Depending on the behavior and orientation of neurons, such behavior may need many computations stages where each stage enables the network to learn by adjusting weights called the training phase. Deep learning emerged as an amazing technology that gave us the potential to enhance learning in every field of human life [15]–[17].

CNN is very similar to other neural nets and is a well-known algorithm in deep learning. In 1980, K. Fukushima proposed the neocognitron [18], which is also regarded as the predecessor of CNN. This discovery was based on the findings of Hubel & Wiesel [19] describing that a layered structure of neurons is used by mammals to visually perceive the world. [20], [21] established the modern framework of CNN.

There are many variants of CNNs as we shall discuss in the following section but the basic architecture of CNN remains the same, broadly consisting of 3 layers –

Aayush Jain is with Department of Information Technology, Delhi Technological University, Delhi-110042, India. (email: aayush89j@gmail.com).

Dinesh Kumar Vishwakarma is with Department of Information Technology, Delhi-110042, India. (email: dvishwakarma@gmail.com, dinesh@dtu.ac.in).

978-1-7281-4988-2/20/$31.00 ©2020 IEEE

IEEE Advancing Technology for Humanity

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:24:13 UTC from IEEE Xplore. Restrictions apply.

convolutional, pooling, and fully-connected layers. Convolution layer consists of many convolution kernels that compute feature maps, the input is convolved using a learned kernel and then non-linear activation function is applied on the convolved output helps in obtaining a new feature map. The pooling layer achieves shift-invariance by dropping the resolution of the feature map, usually sandwiched between two convolution layers. One or more fully connected layers aim to perform high-level reasoning and generate global semantic data. Components of CNN along with their different types are summarized as in Table I.

TABLE I
COMPONENTS OF CNN

Components of CNN	Types
Convolution	Dilated
	Transposed module
	Tiled
	Network in network
	Inception module
Pooling	Mixed
	L_p pooling
	Stochastic
	Multi-Scale orderless
	Spectral
Activation Funtions	Rectified Linear Unit (ReLU)
	Parametric ReLU
	Randomized ReLU
	Leaky ReLU
	Maxout
	Probout
	Exponential linear unit (ELU)

III. VARIANTS OF CNN

In the last decade, we have seen the birth of numerous CNNs. Even the visualization of these models has become extremely challenging because these networks have gotten very deep so lately we stopped keeping track of them and started treating them as black-box models. A plethora of CNN architectures which we see today is a result of various things like improved hardware, ImageNet competition, new ideas, solving specialized tasks, larger datasets, new algorithms and so on. There are various modifications done in existing CNN architecture which improved the network and their performance over time. With numerous variants of CNN, neural networks are slowly becoming for a data scientist or machine learning practitioners what linear regression was one for statisticians. Table II summarizes 10 common variants of CNN used by researchers to optimize the solution for their problem statements. Although there are other variants as well but exploring each and every variant is not done as of now.

IV. CNN FOR VIOLENCE DETECTION

Many variations are proposed to already exist CNN architectures as well as additional deep learning architectures are added to make the classification of violent actions much more accurate.

TABLE II
10 COMMON VARIANTS OF CNN

Name	Year	Parameters	Key Features
LeNet-5	1998	60,000	- Standard template. - Stacking convolutional layer and pooling layers, and ending with one or more fully-connected layers
AlexNet	2012	60 Million	- First to implement ReLU as an activation function
VGG-16	2014	138 Million	- Contributed to the designing of deeper networks
Inception-v1	2014	5 Million	- Rather than stacking convolutional layers, modules/blocks were stacked within which are convolutional layers
Inception-v3	2015	24 Million	- Among the first to use batch normalization
ResNet-50	2015	26 Million	- Though this was not the first to introduce skip connections but popularized skip connections
Xception	2016	23 Million	- CNN based entirely upon depthwise separable convolutional layers.
Inception-v4	2016	43 Million	- Increased model size due to which it worked better than previous versions.
Inception-ResNets	2016	56 Million	- Inception modules converted into Residual Inception Blocks - More Inception modules were added
ResNeXt-50	2017	25 Million	- Scaling up the number of parallel towers within a module

Deep learning architectures, are also based on neural networks, classify violent actions based on the dataset and extracted features using more convolution layers. If we deconstruct a video it is compiled of various frames of images. Video is nothing but an illusion of continuous moving images. If we can classify the image in the frame as either violent image or non-violent image, we can classify the scene of a video as well. Here, the advancement in the convolution neural network algorithm in the proposed frameworks of violence detection is exhaustively discussed. Table III summarizes the recognition methods that use CNN based classification.

V. POPULAR DATASETS USED

Due to the novelty and subjectivity of the concerned research area, the field of violence detection uses a variety of datasets available and does not contain a single main dataset to study. The datasets used for studying the performance of the models proposed by researchers that we explored are summarized in Table IV with their brief overview and characteristics. Most of the datasets use ACC(accuracy), which is the ratio of true results over total cases examined, and AUC(area under the curve), which means that a classifier will assign a higher probability to a positive example than a negative example, for evaluation purpose.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:24:13 UTC from IEEE Xplore. Restrictions apply.

TABLE III
VIOLENCE DETECTION TECHNIQUES USING CNN

Reference	Key Features/Other Characteristics	Advantages	Short Comings	Accuracy
[22]	3D ConvNets for detecting violence without using any prior knowledge, supervised learning for Model training and back-propagation method for computing gradients	- efficient and accurate since directly operate on the image pixel -Automatically learn the video features	- Requires improvement in accuracy - Possibility to detect mid-level concepts	achieve 91.00% accuracy rate on hockey dataset
[23]	2 approaches to use CNNs for classification 1. use in the end-to-end fashion, take raw features as input and at its last layer produce the classification result 2. Use CNNs for feature representations.	Video clips consisting shot, scream and heavy metal music, have high scores because of explicit violent relevance	- audio-level violent identified as non-violent - Non-violent videos with elements like cheer, shutter sound, clapping, and rock music, etc. get high scores.	- Avegage Precision(AP) of 0.485 and 0.291 on MediaEval 2015 dataset on the validation and testing sets for Visual-only + End-to-end CNN based on Audio-only
[24]	- extracting frame-level characteristics from a video using AlexNet model as CNN A variant of the LSTM, that uses convolutional gates, aggregates the frame-level characteristics.	Localized Spatio-temporal features can be captured using CNN + convLSTM helping in the analysis of local motion taking place in the video. - violent video classification by an end-to-end trainable deep NeuralNet	- unable to distinguish violent actions, marks certain videos as non-violent, and unable to outperform the previous technique on Violent-Flows dataset,	- 97.1±0.55% accuracy on hockey dataset - 100±0% accuracy on movies dataset - 94.57±2.34% accuracy on Violent-Flows dataset
[25]	- extracting the convolutional feature maps using 2 Deep CNN models as an extractor	- improved trajectory to capture long-term action information by integrating Deep NN and trajectory - can be expanded to crowd scene videos	- detection of violence regions are yet to be taken into consideration	- approx. 98% on Hockey Fights dataset - approx. 92.5% accuracy on Crowd Violence dataset
[26]	- Pre-trained light-weight MobileNet CNN helps in reducing bulk processing of useless frames. - 3D CNN for feature extraction	- nearest security department or a police station is alerted if violence detected - unnecessary processing of useless frames is reduced - achieved better accuracy	- require good hardware - any resource constraint devices might not be able to deploy the proposed model	- violent crowd dataset - 98% accuracy - movies dataset - 99% accuracy - hockey fight dataset - 96% accuracy
[27]	representative image classified and final decision for the sequence obtained using 2D CNN	- fast enough - real-time application - a good tradeoff between accuracy and computational time.	- better results are achieved only because the camera position was static	- 99_0.5% accuracy on movies dataset - 94.6_0.6% accuracy on hockey dataset - 91.42_0.14% accuracy on Behave dataset
[28]	Convolutional Neural Networks classifier AlexNet based CNN with a subset of ImageNet classes	- classification Of Large scale sports video - existing information or pre-processing not required	- High computation - low accuracy level	- 65.4% accuracy achieved, by Fine-tuning top 3 layers, on UCF-101 dataset
[29]	- 2-stream CNN for extracting characteristics on motion optical flows and static frames - longer-term temporal dynamics captured using LSTM applied on top of the 2-stream CNN characteristics	- very effective model to detect violence	- additional computational complexity	- Mean AP - 0.296 in the violence detection Subtask - In the induced effect detection subtask, accuracy of 0.418 and 0.488 for arousal and valence respectively
[30]	ScatterNet used for human pose estimation	- needs fewer learning examples - real-time processing is done using cloud	- vast amounts of unlimited computational power and resources are required	- 87.6% accuracy. on the Aerial Violent Individual (AVI) dataset
[31]	Convolutional Neural Networks classifier	- needs lesser computational resources - the model can be trained on CPU only	- model not tested on the popular benchmark datasets	- 95% accuracy on [31] dataset
[32]	Simultaneous crowd counting, violence detection, and crowd density level classification using ResnetCrowd	- individual task performance boosted by multi-task approach (most notably for violent behavior detection)	- unable to properly investigate - lack of an appropriately labeled multi-task dataset	- 9% boost in ROC curve AUC on [32] dataset
[33]	- describing spatial information in a video frame using Deep CNN (ResNet50 / VGG19 / Xception). - bi-directional LSTM to learn a video-level classifier from both violent/non-violent video sequences	- very effective model to detect violence	- very huge training dataset for violence videos required	using ResNet50 + LSTM :- - 75.73% accuracy on Real-Violent Dataset - 88.74% accuracy on Movie Dataset - 83.19% accuracy on Hockey Fight Dataset
[34]	- SELayer-3D CNN	- Higher detection accuracy and efficiency.	- reduced accuracy if lower proportion of violence	- accuracy of 98.08% on the CrowdViolence

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:24:13 UTC from IEEE Xplore. Restrictions apply.

		- fast real-time processing - processing speed of 500 fps in real-time	samples in the data set - require big enough training dataset for violence videos	dataset - 99.0% accuracy on the HockeyFight dataset
[35]	2 DNNs frameworks - 3D-based convolutional neural network + CNN-LSTM	- incorporates more specialized concepts for classifying violent scene during violence detection.	- lesser accuracy than other accurate and highly effective models	- MediaEval-2013-VSD data set - 63% accuracy
[36]	CNN + LSTM (as spatial feature extractor and temporal relation learning method respectively)	- 4X faster than the fastest known model - speed of the base model is 131fps	- the model doesn't perform to the fullest due to lesser availability of data	- accuracy of 98 % on the Hockey Fight Dataset
[37]	Spatiotemporal Encoder + Bidirectional ConvLSTM architecture.	- better performance for complex and more heterogeneous datasets like the Violent Flows dataset	- the method is not consistently best, require more training data for better and accurate results	- Hockey Fights dataset - 96.96±1.08% accuracy - Movies dataset - 100±0% accuracy - Violent Flows dataset - 92.18±3.29%. accuracy
[38]	pre-trained deep learning model VGG-Net	---	- Needs more accuracy improvements	- accuracy of 75.00% on hockey fight dataset

TABLE IV
SUMMARY OF POPULAR DATASETS USED FOR VIOLENCE DETECTION

S. No.	Dataset Name	Year	Description	Evaluation Metrics
1	Caviar	2004	- 28 sequences of videos having 26500 labeled frames - 6 groups of different activity scenario.	---
2	Behave	2006	- Only clip 1 annotated (52:11 minutes) - A fixed point of view Simulated actions - 200,000 frames	ACC AUC
3	Souza	2010	- 400 videos (50% fight scenes)	ACC
4	Giannakopoulos	2010	- 50 clips from 10 movies (2.5 hours total)	Recall F1 Precision
5	Movies	2011	- 200 video clips from action movies	ACC AUC
6	Hockey Fight	2011	- 1000 clips - 50 frames per clip	ACC AUC
7	Perperis	2011	- 25 movie segments (aprox. 1 min each)	ACC Recall F1 Precision
8	UFC 101	2012	- 13,000 clips	ACC
9	Violent Flows	2012	- 246 videos from YouTube (1-7 seconds)	ACC AUC
10	SBU Kinect Interactions	2012	- 21 sets of 8 violent interactions - total of 300 interactions approx.	ACC
11	Violent Scene Dataset	2014	-Train: 25 Hollywood movies, 32678 shots -Test: 7 Hollywood movies, 11245 shots - 86 web videos	AUC
12	LIRIS-ACCEDE (medieval)	2015	- 10900 short video clips (8-12 seconds) -Train:100 movies, 6144 shots - Test:99 movies, 4756 shots	AUC
13	RE-DID(real-life events-Dyadic Interactions Dataset)	2015	- Real Life Scenarios - urban fights - 30 videos Length: 0:20 to 4:20	ACC
14	KARD	2017	- 18 activities - divided in 10 gestures and 8 actions - 540 sequences having resolution of 640 × 480 pixels at 30 fps - total of 1hr of videos	ACC Precision Recall
15	Crowd-11	2017	- 6272 videos - 3005 scenes - 621,196 frames	ACC

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:24:13 UTC from IEEE Xplore. Restrictions apply.

VI. Conclusion

With the rising population and increasing need for surveillance has posed growing demands of systems that are capable to detect violent acts automatically. Violence detection is already an interesting research field and CNN's have made an exceptional breakthrough in the field of detecting violence. The objective of this review is to present state-of-the-art research in the field of violence detection specifically using CNN. Firstly, a brief introduction to CNN and its component is presented, followed by the various variants of CNN. We have presented thorough research on the various CNNs used for violence detection in videos, their advantages, and shortcomings. In addition, we have surveyed major publically available popular datasets used to test the proposed models along with their characteristics and the evaluation metrics used on those datasets. Although, a great deal of success has been achieved by CNN's yet there a lot of issues that require further investigation

References

[1] W. Sultani, C. Chen, and M. Shah, "Real-World Anomaly Detection in Surveillance Videos," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 6479–6488, 2018.

[2] H. U. Keval, "Effective design, configuration, and use of digital CCTV," *Crit. Rev.*, no. April, pp. 1–289, 2009.

[3] https://vidimensions.com/, "Vi Dimensions | We Discover. Going Beyond Detection." [Online]. Available: https://vidimensions.com/. [Accessed: 10-Dec-2019].

[4] I. S. Gracia, O. D. Suarez, G. B. Garcia, and T. K. Kim, "Fast fight detection," *PLoS One*, vol. 10, no. 4, pp. 1–19, 2015.

[5] O. Deniz, I. Serrano, G. Bueno, and T. K. Kim, "Fast violence detection in video," *VISAPP 2014 - Proc. 9th Int. Conf. Comput. Vis. Theory Appl.*, vol. 2, pp. 478–485, 2014.

[6] D. K. Vishwakarma, R. Kapoor, and A. Dhiman, "A proposed unified framework for the recognition of human activity by exploiting the characteristics of action dynamics," *Rob. Auton. Syst.*, vol. 77, pp. 25–38, 2016.

[7] D. K. Vishwakarma and K. Singh, "Human activity recognition based on spatial distribution of gradients at sublevels of average energy silhouette images," *IEEE Trans. Cogn. Dev. Syst.*, vol. 9, no. 4, pp. 316–327, 2017.

[8] D. K. Vishwakarma, R. Kapoor, and A. Dhiman, "Unified framework for human activity recognition: An approach using spatial edge distribution and ℜ-transform," *AEU - Int. J. Electron. Commun.*, vol. 70, no. 3, pp. 341–353, 2016.

[9] L. Tian, H. Wang, Y. Zhou, and C. Peng, "Video big data in smart city: Background construction and optimization for surveillance video processing," *Futur. Gener. Comput. Syst.*, vol. 86, pp. 1371–1382, 2018.

[10] P. Zhou, Q. Ding, H. Luo, and X. Hou, "Violent Interaction Detection in Video Based on Deep Learning," *J. Phys. Conf. Ser.*, vol. 844, no. 1, 2017.

[11] C. Dhiman and D. K. Vishwakarma, "A review of state-of-the-art techniques for abnormal human activity recognition," *Eng. Appl. Artif. Intell.*, vol. 77, no. August 2018, pp. 21–45, 2019.

[12] M. Ramzan *et al.*, "A Review on State-of-the-Art Violence Detection Techniques," *IEEE Access*, vol. 7, pp. 107560–107575, 2019.

[13] P. D. Garje, M. S. Nagmode, and K. C. Davakhar, "Optical Flow Based Violence Detection in Video Surveillance," *2018 Int. Conf. Adv. Commun. Comput. Technol. ICACCT 2018*, pp. 208–212, 2018.

[14] Z. Guo, F. Wu, H. Chen, J. Yuan, and C. Cai, "Pedestrian violence detection based on optical flow energy characteristics," *2017 4th Int. Conf. Syst. Informatics, ICSAI 2017*, vol. 2018-Janua, no. Icsai, pp. 1261–1265, 2017.

[15] J. Schmidhuber, "Deep Learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, 2015.

[16] J. Gu *et al.*, "Recent advances in convolutional neural networks," *Pattern Recognit.*, vol. 77, pp. 354–377, 2018.

[17] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," 2015.

[18] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biol. Cybern.*, pp. 193–202, 1980.

[19] T. N. Wiesel and H. Hubel, "Receptive fields and functional architecture of monkey striate cortex," *J. Physiol.*, pp. 215–243, 1968.

[20] R. E. Howard, W. Hubbard, and L. D. Jackel, "Handwritten Digit Recognition with a Back-Propagation Network," *Adv. Neural Inf. Process. Syst.*, pp. 396–404, 1990.

[21] LeCun, Yann, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, pp. 2278–2324, 1988.

[22] C. Ding, S. Fan, M. Zhu, W. Feng, and B. Jia, "Violence detection in video by using 3D convolutional neural networks," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8888, pp. 551–558, 2014.

[23] G. Mu, H. Cao, and Q. Jin, "Violent Scene Detection Using Convolutional Neural Networks and Deep Audio Features," 2008, pp. 645–651.

[24] S. Sudhakaran and O. Lanz, "Learning to detect violent videos using convolutional long short-term memory," *2017 14th IEEE Int. Conf. Adv. Video Signal Based Surveillance, AVSS 2017*, 2017.

[25] P. Yun, J. Jiao, and M. L. B, "Trajectory-Pooled Deep Convolutional Networks for Violence Detection in Videos," *Comput. Vis.*, vol. 2, no. 6140021318, pp. 3–15, 2017.

[26] F. U. M. Ullah, A. Ullah, K. Muhammad, I. U. Haq, and S. W. Baik, "Violence detection using spatiotemporal features with 3D convolutional neural network," *Sensors (Switzerland)*, vol. 19, no. 11, pp. 1–15, 2019.

[27] I. Serrano, O. Deniz, J. L. Espinosa-Aranda, and G. Bueno, "Fight Recognition in Video Using Hough Forests and 2D Convolutional Neural Network," *IEEE Trans. Image Process.*, vol. 27, no. 10, pp. 4787–4797, 2018.

[28] A. Karpathy and T. Leung, "Large-scale Video Classification with Convolutional Neural Networks," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 10–20, 2018.

[29] Q. Dai *et al.*, "Fudan-Huawei at MediaEval 2015: Detecting violent scenes and affective impact in movies with deep learning," *CEUR Workshop Proc.*, vol. 1436, pp. 5–7, 2015.

[30] A. Singh, D. Patil, and S. N. Omkar, "Eye in the sky: Real-time drone surveillance system (DSS) for violent individuals identification using scatternet hybrid deep learning network," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 2018-June, pp. 1710–1718, 2018.

[31] G. Sakthivinayagam, R. Easawarakumar, A. Arunachalam, and M. Pandi, "Violence Detection System using Convolution Neural Network," *SSRG Int. J. Electron. Commun. Eng.*, vol. 6, pp. 6–9, 2019.

[32] M. Marsden, K. McGuinness, S. Little, and N. E. O'Connor, "ResnetCrowd: A residual deep learning architecture for crowd counting, violent behaviour detection and crowd density level classification," *2017 14th IEEE Int. Conf. Adv. Video Signal Based Surveillance, AVSS 2017*, 2017.

[33] E. Ditsanthia, L. Pipanmaekaporn, and S. Kamonsantiroj, "Video Representation Learning for CCTV-Based Violence Detection," *TIMES-iCON 2018 - 3rd Technol. Innov. Manag. Eng. Sci. Int. Conf.*, pp. 1–5, 2019.

[34] B. Jiang, F. Xu, W. Tu, and C. Yang, "Channel-wise Attention in 3D Convolutional Networks for Violence Detection," *Proc. - 2019 Int. Conf. Intell. Comput. Its Emerg. Appl. ICEA 2019*, pp. 59–64, 2019.

[35] B. Peixoto, B. Lavi, and P. Martin, "Toward subjective violence detection in videos," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, pp. 8276–8280, 2019.

[36] A. M. R. Abdali and R. F. Al-Tuma, "Robust Real-Time Violence Detection in Video Using CNN and LSTM," *SCCS 2019 - 2019 2nd Sci. Conf. Comput. Sci.*, pp. 104–108, 2019.

[37] A. Hanson, K. Pnvr, S. Krishnagopal, and L. Davis, "Bidirectional convolutional LSTM for the detection of violence in videos," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11130 LNCS, pp. 280–295, 2019.

[38] S. Mukherjee, R. Saini, P. Kumar, P. P. Roy, D. P. Dogra, and B.-G. Kim, "Fight Detection in Hockey Videos using Deep Network," *J. Multimed. Inf. Syst.*, vol. 4, no. 4, pp. 225–232, 2017.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:24:13 UTC from IEEE Xplore. Restrictions apply.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/344026575

Statistical Analysis of Relation Between Developmental Status of Nations and their Banks' NPA using Hypothesis Testing

Research · August 2020

DOI: 10.34218/IJM.11.8.2020.082

CITATIONS
0

READS
17

6 authors, including:



Mridul Gupta
Delhi Technological University
1 PUBLICATION **0** CITATIONS
SEE PROFILE



Nand Kumar
Delhi Technological University
47 PUBLICATIONS **75** CITATIONS
SEE PROFILE



Nimish Sharma
Delhi Technological University
1 PUBLICATION **0** CITATIONS
SEE PROFILE

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

Efficiency analysis banks in India. View project

Dates Extended for Full manuscript Submission at ICAPIE 2016, Delhi, India View project

All content following this page was uploaded by Mridul Gupta on 01 September 2020.

The user has requested enhancement of the downloaded file.

International Journal of Management (IJM)
Volume 11, Issue 8, August 2020, pp. 910-922, Article ID: IJM_11_08_082
Available online at http://www.iaeme.com/IJM/issues.asp?JType=IJM&VType=11&IType=8
ISSN Print: 0976-6502 and ISSN Online: 0976-6510
DOI: 10.34218/IJM.11.8.2020.082

© **IAEME** Publication **Scopus** Indexed

STATISTICAL ANALYSIS OF THE RELATION BETWEEN THE DEVELOPMENTAL STATUS OF NATIONS AND THEIR BANKS' NPA USING HYPOTHESIS TESTING

Gangesh Chawla
Department of Production & Industrial Engineering, Delhi Technological University, Bawana Rd, Shahbad Daulatpur Village, Rohini, Delhi, India

Mridul Gupta*
Department of Mechanical Engineering, Delhi Technological University, Bawana Rd, Shahbad Daulatpur Village, Rohini, Delhi, India

Nimish Sharma
Department of Mechanical Engineering, Delhi Technological University, Bawana Rd, Shahbad Daulatpur Village, Rohini, Delhi, India

Keshav Aggarwal
Department of Production & Industrial Engineering, Delhi Technological University, Bawana Rd, Shahbad Daulatpur Village, Rohini, Delhi, India

Ranganath M. Singari
Department of Production & Industrial Engineering, Delhi Technological University, Bawana Rd, Shahbad Daulatpur Village, Rohini, Delhi, India

Nand Kumar
Department of Humanities, Delhi Technological University, Bawana Rd, Shahbad Daulatpur Village, Rohini, Delhi, India
*Corresponding Author

ABSTRACT

HDI is the most common index which is used in measuring the development of a nation. Life expectancy, literacy rate and per capita income are the main components of Human Development Index (HDI). Health, standard of living and education are the main aspects it covers. The subsequent topic this paper covers is Non-performing Assets (NPA) or Bad Loans. These NPAs are caused due to an improper structure of lending and a lack of application of suitable technology in the process. The loan

taking capability of the masses is influenced by another factor called Economic Inequality, which is the difference between individuals or populations in terms of income, wealth and assets. We will observe the effect of health, education and standard of living of a country on NPAs of Banks of that country which in turn affects the interest rates levied on loans, thereby swaying the loan taking capacity of the people. In our research we analyse the NPAs of top 2 banks of 15 developing nations and 14 developed nations and formulate a Statistical Model to determine its relation with their respective HDI. We make use of T-Test and a statistical approach to develop the model and come up with a suitable conclusion

JEL Codes: C12, E10, G21, O11

Keywords: Non-performing Assets (NPA), Developing Nations, Developed Nations, t-test, Human Development Index (HDI)

Cite this Article: Gangesh Chawla, Mridul Gupta, Nimish Sharma, Keshav Aggarwal, Ranganath M. Singari and Nand Kumar, Statistical Analysis of the Relation between the Developmental Status of Nations and their Banks' NPA using Hypothesis Testing, *International Journal of Management,* 11(8), 2020, pp. 910-922.
http://www.iaeme.com/IJM/issues.asp?JType=IJM&VType=11&IType=8

1. INTRODUCTION

Finance, which is the study of the flow of money, has become a fundamental aspect of our day to day lives. Without involving finance, we cannot even secure our future. Finance is involved in operations as simple as depositing money into a savings account to complex operations like making the budget of a country. The amount along with the flow of money in the market is very essential for a country's economy, which in-turn is vital for the country's growth.

A major aspect of cash flow in the markets are Financial Instruments, which can be defined as the proof of ownership of an asset in the form of stocks, bonds or shares. These financial instruments are bought by using Investment Vehicles i.e. products used by investors to invest in these stocks, shares, bonds, etc. and gain profits.

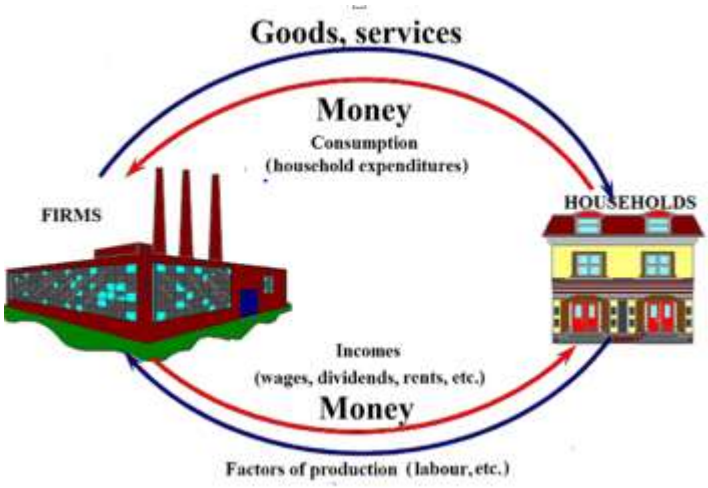


Figure 1 Flow of cash into the economy

The investments can be of various types [2]:

Ownership Investments; these are the investments which are made in order to own a portion or whole of something. These involve investment vehicles like Stocks, Real Estates, precious metals and objects, etc.

Lending Investments; these are the investments in which the investor lends money to another person or entity. These involve investment vehicles such as Bonds, Certificates of deposits, etc.

Pooled Investment; this is a type of investment in which the investors invest their money not separately but by pooling their money. These involve investment vehicles such as Mutual Funds.

Since people are now becoming more aware and have started taking interest in these investment options, the cash flow and thus the economy has boosted. There are various companies which are not conventional banks but still help people invest their money. These companies are called Non-Banking Financial Company or NBFC. A NBFC is an organization enlisted under the Companies Act (1956) which does not have a full banking license but can be involved in the act of lending, hire purchase, leasing, insurance, chit funds, stocks, share acquisition and in some cases receiving deposits.

Another financial institution which helps the investors to invest is a Bank. The banks are the oldest and the most conventional form of financial institution and are the most trusted method of investing. Since the beginning, the banks have attracted a lot of investments and were able to handle them successfully, but this has not been the scenario since the last decade. Earlier, as technology evolved the banking industry was very good in integrating that technology in order to provide a better customer experience, moreover the number of investors were also less. This helped banks handle all the investments easily and remain profitable.

Now, the banks are not able to handle the investments and loans, and are thus facing losses. All this happened after the Financial Crisis of 2008, during which the banks were very busy handling the new rules, regulatory requirements and fines imposed on them, thus making innovation a very distant priority. Due to lack of technology, the Indian banks especially were unable to keep their customers happy in terms of user experience and ease of access. All this eventually led to the shift of the investors towards new and high-tech NBFCs. This in-turn led to losses and the banks started merging themselves in order to remain functioning.

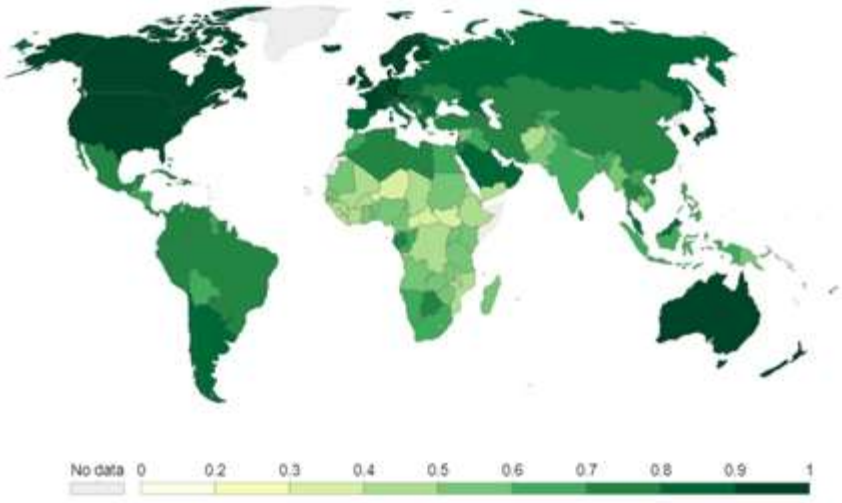


Figure 2 Worldwide Human Development Index [3]

All these losses and merging/closing of the banks has directly affected the customers who take loans from these banks where the interest rates are skyrocketing nowadays. Loans are a crucial part of today's economy as they help people reach personal endeavors or help finance start-ups. Higher interest rates on the loans and lack of application of suitable technology has increased the number of NPAs or Bad Loans beyond the critical levels. An NPA/Bad Loans can be defined as the loans issued by the banks whose one or more installments are due for a period greater than 90 days.

These NPAs can also be considered a factor in the categorization of the countries into developed and developing as according to our hypothesis the average NPA of a developed country will be less than that of a developing country. According to the United Nations the division of the countries into developed and developing is based on the HDI which is the Human Development Index defined as the statistical measure formulated to measure a country's social and economic development. It covers areas such as life expectancy, expected year of schooling and per capita income of a country.

During our research we chose 15 developing and 14 developed countries according to a World Bank report [4] and selected 2 banks of each country that we'll be examining. We analyzed the annual report and investors presentations of these banks from 2014-2018 and found out the NPA ratio released by the banks for each year.

2. STATEMENT OF PROBLEM

Modern day countries are categorized by the World Bank as developing and developed on the basis of Gross National Income (GNI) per capita. The country whose GNI per capita falls under lower income or lower middle-income range is classified as developing while the one who's GNI per capita falls under upper middle income or higher income is classified as developed nation. Besides this, a country's development status is also determined by an important aspect, namely Human Development Index (HDI). The HDI factors in 3 broad indices viz health, standard of living and education, giving equal weightage to each. Although Gross Domestic Product (GDP) has become the standard measure to assess the general economic performance of a country, GDP alone couldn't be the measure to determine the development status of a country as the assumption of using per capita GDP as a way to measure the development is flawed in itself [5]. In the case of India, non-deprivation in income does not necessarily mean non-deprivations in health and education [6].

In this study, we have established an assumption that high NPA ratios of the Banks of a country are an indirect effect of its low HDI. Some probable reasons for a high NPA ratio due to low HDI have been discussed further in this section.

A deficient standard of living is a result of low income. Lower incomes of the common masses can be a reason for the inability of loan repayment to Banks, resulting in a rise in NPA ratios. Furthermore, degraded health of an individual doesn't comply with efficient functionality and hence it can also become a factor in incapability of performing work. With no work and hence no income generation, loan repayment may become difficult, making Banks witness a high NPA ratio. Also, lack of decent and sufficient education level begets lack of good-pay jobs, thereby generating insufficient income to repay loans and spiking NPA rates of Banks and that unemployment rate affects impaired loans positively [7].

Exploring another perspective; at times the country Governments launch certain schemes to advance financial inclusion in their countries and promote the formal lending system against the informal lending system. This results in a higher number of loans being lent out to the customers. But with a low HDI and low-income level, repayment of these loans is not up to the mark, leading to escalated NPA ratios than usual. For instance, in India, a Right to

Information (RTI) report revealed that NPAs against loans under Pradhan Mantri MUDRA Yojana Scheme (PMMYS) were on a rise. As much as INR 1.25 bn was waived off by 2 Indian public-sector Banks between 2015-19 under the program [8]. The total NPA contribution from MUDRA loans did not exceed 2.68% of total NPAs but if the trend continued it would be one of the key contributors for the NPAs. A 34% increase in NPA in only one state like Gujarat in the past 1 year can be a major setback for Indian Economy (Nagarkar, Bhatt and Nagar, 2020) [9].

Further, the paper follows up with the verification of our assumption through a statistical model.

3. LITERATURE REVIEW

Saibal Ghosh [10] examines the relation between corporate leverage and bank's NPAs using the data obtained for the manufacturing sector in India for 1993-2004. The study involves application of debt to equity ratio which is the most common measure of corporate leverage for the measurement of the leverage.

Mandira Sarma [11] has formulated a model using the index of financial inclusion to identify the factors significantly affecting the financial inclusion. Using this model, they present a cross country empirical analysis of the relationship between financial inclusion and development.

Namita Rajput [12] uses the secondary data collected from reports, journals and websites to know about the Indian banking sector to highlight the composition, development and management of non-performing assets.

Suman Seth [13] discusses the measurement of Poverty and Human Development using the United Nations Development Program's (UNDP) Global Human Development Report. Focusing on the HDI and the Poverty indices outlines the methodological evolution of different indices over the last 2 decades. Critically evaluates the measures and discusses the possible improvements that can be made.

Piet H. Renooy [14] focusses on the extent the informal economy may be considered a specifically urban phenomenon. Informal economy is divided into three types: the black, the communal and the household economy and economy indications to find these respectively.

Jeevan Nagarkar [9] has connected the issue of NPA with the MUDRA Yojna scheme as the increase in NPA has been adversely affecting the growth of the GDP as well as the stability of the banking sector of the country. He has also discussed the effect of MUDRA Yojna on self-employment and wealth generation.

Salini R Chandran [15] has identified trends in NPA in the Indian Banking system and its impact on Asset Quality. Through the bankers' perspective, the reasons for the NPA are identified and some measures are provided to minimize the NPAs. An effective use of major analysis tools like percentages, growth rates, standard deviation, mean, chi-square test and Z-test has been applied over primary and secondary data which was collected over a period from 2002 to 2017.

Nammi Lalitha [16] offers a comparative study of NPAs over Public sector banks and Private sector banks and its impact. It studies the enormity of NPAs, its causes, trends, the strategies and recovery measures followed by banks to counter it. The author also presents several suggestions to the private and public sector banks to tackle the problems posed by NPAs.

Raj Kishor Pradhan [17] has investigated the negative impact of the NPAs on the Indian Economy. The paper discusses the menace of NPAs in the Indian banking sector and the challenges offered by it. Various possible solutions have also been explored by him.

P.D. Jeromi [18] has presented his study on regulations of informal financial institutions with a special focus on money lenders in Kerala, India. Workings of the informal financial sector has been investigated and special emphasis has been given on the unruly and unlawful practices pursued by them which pose a threat to the stability of the formal economy. The paper has found incompetence in the present supervisory and regulatory mechanisms to deal with the informal financial institutions and has provided some suggestions to counter it.

Tae Kyun Kim [19] intends to give an introduction to the T-test as a parametric statistic where the probability distribution of the population is known which a fixed has set of parameters. It is a hypothesis testing procedure based on certain assumptions and preconditions. This paper describes the situations where T-test can be applied and the procedure to apply it.

Ahlem Selma Messai [7] has ascertained some micro and macro determinants of Non-Performing Loans or NPA in a sample of European Banks. The banks were put under study and observed for the period of 2004 to 2008 to establish a relation between NPAs and factors like GDP growth, unemployment rate, and return on assets, interest rates and loan loss reserves to total loans ratio.

4. RESEARCH METHODOLOGY

Step 1: Choosing the right dataset and statistical test

It was kept into consideration to choose an extensive dataset while conducting the study. Countries from each continent were selected to bring in a diverse and a global perspective to the picture. Special care was taken to select only commercial banks and not the central banks as the latter doesn't have a direct relation with the people. Also, the Net NPA ratios of the banks were considered for the study as it does not factor in provisions set aside by banks. On the basis of data available, the researchers chose the Student's T-test for hypothesis testing instead of Z-test because the variance of the population is not known.

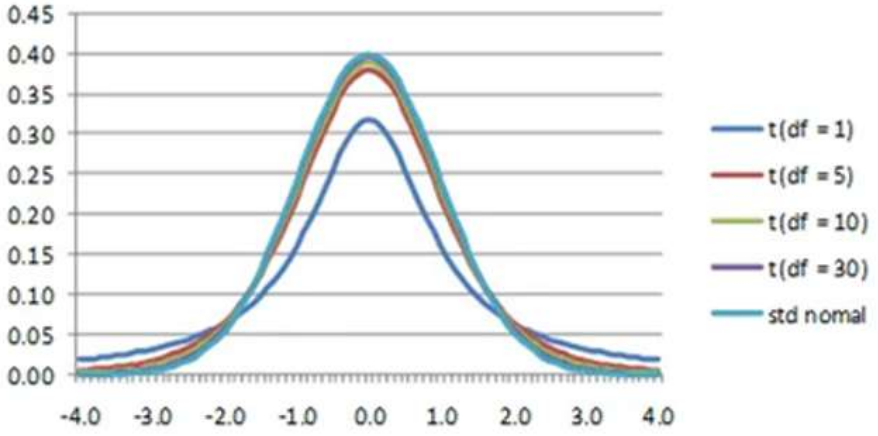


Figure 3 Student's T-Distribution[21]

Student's T-Test: The Student's T-Distribution was first published by William Seally Gosset in his 1908 paper which was named as Biometrika[20]. The T-test is a type of statistical test which is used in comparison of means of two groups. It can broadly be of 2 types, Independent t-test and Paired t-test [19]. Independent t-test is used to compare 2 groups

which are not dependent on each other. At the beginning of the test a Null Hypothesis and an Alternative Hypothesis are defined. The end result of the Student's T-test tells us whether to reject H_o or not,

The graph here gives us very important information regarding the Student's t-test which is if the degree of freedom (df) increases i.e. it reaches 30, the graph of t-distribution tends to that of normal distribution.

Step 2: Setting up the Hypothesis

Once the Student's T-test was chosen, the researchers decided the null and the alternative hypothesis:

H_0: NPAs of developed countries - NPA of developing countries > 0.

H_1: NPAs of developed countries - NPA of developing countries ≤ 0.

Step 3: Setting the Assumptions

The researchers have made some assumptions for the performing the hypothesis testing using the Student's T-test which are as follows:

1. The data sample we have chosen from the population is normally distributed as validated by the Central Limit Theorem.
2. The population variances of developed and developing countries are unknown but equal and the two populations are independent of each other.

Step 4: Calculating the Weighted Mean of NPA Percentage for every Bank

Once the NPA percentage for all the countries for 5 years (2014-18) is collected, the weighted mean is calculated using the formula:

$$weighted\ \underline{x} = \frac{\sum_{i=1}^{n} (x_i \times w_i)}{\sum_{i=1}^{n} w_i}$$

where n=5, x_1=0.2, x_2=0.4, x_3=0.6, x_4=0.8, x_5=1 and w_i stands for NPA percentages for the years 2014, 2015, 2016, 2017, 2018 respectively.

Step 5: Calculating the Mean NPA percentage for Nations

Firstly, the weighted average of 2 banks of every country is further averaged to get the Mean NPA percentage for each country. Then the 15 means we get are averaged to get the mean NPA percentage for Developed and Developing countries.

Step 6: Calculating the Pooled Variance and Standard Deviation

Firstly, the variance for both developing and developed nations is calculated using the formula:

$$\sigma^2 = \frac{\Sigma (x_i - \underline{x})^2}{n-1}$$

Where x_i is average NPA percentage of each country, $\underline{x}$ is the mean NPA percentage for developed or developing countries and n is the number of developed or developing countries.

Once the variance for both developed (σ_x^2) and developing countries (σ_y^2) is calculated, the Pooled Variance (s_p^2) and Pooled Standard Deviation (s_p) is calculated as follows:

$$s_p^2 = \frac{(n_x - 1)\sigma_x^2 + (n_y - 1)\sigma_y^2}{n_x + n_y - 2}$$

$$s_p = \sqrt{\frac{s_p^2}{n_x} + \frac{s_p^2}{n_y}}$$

Where n_x is the number of developed countries i.e. 14 and n_y is the number of developing countries i.e. 15

Step 7: Calculating the T Value

After the pooled variance and standard deviation are calculated, the T value is calculated for which we require:

$\underline{x}_x$= Mean NPA Percentage of Developed Nations

$\underline{x}_y$= Mean NPA Percentage of Developing Nations

$$\underline{d} = \underline{x}_x - \underline{x}_y$$

μ_0= Hypothesized Mean

and Pooled Standard Deviation(s_p)

$$T = \frac{\underline{d} - \mu_0}{s_p}$$

Step 8: Deciding the Confidence Interval

Once the T value is calculated, we use the t-table [22] to check for the appropriate confidence interval. Using the

$$\text{Degree of Freedom}(\nu) = n_x + n_y - 2$$

we look for the nearest value to the calculated T-value and if the T-value is more than the t-value located from the table, the null hypothesis is disregarded, else the alternative hypothesis is disregarded.

4.1. Numerical Illustration

NPA Percentage: (In accordance with Step 1)

Table 1 Developed

S. NO.	Continent	Country	Bank Name	2018	2017	2016	2015	2014
1	Asia	Japan	Mitsubishi UFJ Financial Group	0.62	0.88	1.11	1.19	1.16
2			Mizuho Financial Group	0.62	0.66	1	1	1.2
3	North America	USA	JPMorgan Chase	0.52	0.69	0.84	0.84	1.05
4			Bank of America	0.22	0.27	0.38	0.28	2.22
5	Asia	Singapore	DBS Group	1.5	1.7	1.4	0.9	0.9
6			United Overseas Bank	1.5	1.8	1.5	1.4	1.2
7	Europe	Norway	Bank Norwegian	7.6	5.7	4.2	4	2.8
8			Storebrand Bank	0.4	1	0.7	0.6	0.5
9	Australia	Australia	National Australia Bank	0.71	0.7	0.85	0.63	1.47
10			Westpac	0.53	0.49	0.55	0.45	0.53
11	Europe	Britain	HSBC Holdings	0.2	0.27	0.2	0.19	0.14
12			Lloyds Banking Group	1.9	1.6	1.8	2.1	2.9
13	Europe	Germany	Deutsche Bank	0.8	0.76	0.8	0.67	0.61
14			DZ Bank	1.5	1.7	1.5	1.7	1.9

S. NO.	Continent	Country	Bank Name	2018	2017	2016	2015	2014
15	North America	Canada	Toronto- Dominion Bank	1	0.9	1	1.1	1.1
16			Bank of Nova Scotia	0.6	0.43	0.49	0.44	0.46
17	Europe	France	Crédit Agricole	11.5	12.9	13.9	13.3	11.6
18			Societe Generale	3.6	4.4	5	5.3	6
19	Asia	Israel	Bank Hapoalim	0.76	0.77	1.26	1.75	
20			Bank Leumi	0.89	1.01	1.2	1.3	1.7
21	Asia	South Korea	KB Financial Group	0.6	0.7	0.9	0.84	1.03
22			Shinhan Financial Group	0.53	0.62	0.74	0.87	1.15
23	Europe	Italy	Unicredit S.P.A	3.2	4.8	7.5	12.2	11.7
24			Intesa Sanpaolo	4.2	5.6	8.2	9.5	9.8
25	Oceania	New Zealand	ANZ Bank New Zealand	0.44	0.42	0.48	0.46	0.89
26			Kiwi Bank	0.06	0.07	0.09	0.15	0.3
27	South America	Uruguay	Bladex	1.12	1.07	1.09	0.78	0.06
28			Banco Itaú Uruguay	1.88	1.8	2.1	1.9	1.7

Table 2 Developing

S. NO.	Continent	Country	Bank Name	2018	2017	2016	2015	2014
1	Asia	India	State Bank of India	10.91	6.9	6.5	4.25	4.95
2			HDFC	1.3	1.05	0.94	0.9	1
3	Africa	South Africa	Absa Group Limited	3.75	3.73	3.84	3.94	4.6
4			Firstrand Bank	2.47	2.27	2.43	2.17	2.32
5	Africa	Nigeria	Zenith Bank	4.9	4.3	3.02	2.2	1.8
6			Access Bank	2.5	4.8	2.1	1.7	2.2
7	Africa	Zimbabwe	CBZ Bank Zimbabwe	16.4	10.9	6.6	6.9	7.39
8			ZB Bank Limited Zimbabwe	8.39	7.08	7.87	10.87	15.91
9	Asia	Malaysia	Maybank Malaysia	1.9	2.1	3.2	2.5	2.2
10			CIMB Bank Malaysia	1.3	3.4	3.3	3	3.1
11	Asia	China	Industrial and commercial bank of china	1.52	1.55	1.62	1.5	1.13
12			China Construction Bank Corporation	1.46	1.49	1.52	1.58	1.19
13	South America	Chile	Banco Santander Chile	2.1	2.28	2.07	2.54	2.8
14			Banco de Credito e Inversiones Chile	1.36	1.4	1.42	1.48	2.29
15	South America	Brazil	Itaú Unibanco Holding	2.3	2.7	2.6	2.8	2.6
16			Brazilian Development Bank	2.95	2.08	2.43	0.02	0.01
17	South America	Argentina	Banco Patagonia	1.7	1.3	1	1.2	1.8
18			Grupo Financiero Galicia	2.47	2.24	2.43	2.46	2.61

S. NO.	Continent	Country	Bank Name	2018	2017	2016	2015	2014
19	Asia	Russia	Sberbank Russia	3.8	4.2	4.4	5	3.2
20			VTB Bank Russia	5.7	5.7	6.4	6.3	5.8
21	Africa	Egypt	Commercial International Bank Egypt	4.06	6.95	6.76	3.98	4.71
22			QNB Alahli Bank Egypt	2.21	2.17	2.54	2.6	2.96
23	Europe	Poland	Powszechna Kasa Oszczędności Bank Poland	4.9	5.5	5.9	6.6	6.9
24			Santander Bank Poland	4.1	5.77	6.58	7.26	8.37
25	Europe	Croatia	Addiko Bank Croatia	7.7	11.6	14.8	11.25	12.2
26			Erste & Steiermärkische Bank	3.2	4	4.9	7.1	8.5
27	North America	Mexico	BBVA Bancomer Mexico	2.3	2.1	2.2	2.4	2.7
28			Banorte Bank Mexico	1.68	1.99	1.79	2.3	2.9
29	South America	Colombia	Bancolombia S.A.	3.1	2.9	2	1.8	1.6
30			Banco de Bogotá Colombia	2.8	2.4	1.7	1.5	1.7

Mean NPA Percentage (Country wise): (In accordance with Step 4 and 5)

Table 3 Developed

S. NO.	Continent	Country	Bank Name	Weighted Mean NPA Percentage	Mean NPA Percentage
1	Asia	Japan	Mitsubishi UFJ Financial Group	0.5396	0.51
2			Mizuho Financial Group	0.4776	
3	North America	USA	JPMorgan Chase	0.4244	0.33
4			Bank of America	0.244	
5	Asia	Singapore	DBS Group	0.848	0.89
6			United Overseas Bank	0.928	
7	Europe	Norway	Bank Norwegian	3.368	1.88
8			Storebrand Bank	0.392	
9	Australia	Australia	National Australia Bank	0.4652	0.39
10			Westpac	0.3076	
11	Europe	Britain	HSBC Holdings	0.128	0.63
12			Lloyds Banking Group	1.136	
13	Europe	Germany	Deutsche Bank	0.4556	0.71
14			DZ Bank	0.964	
15	North America	Canada	Toronto- Dominion Bank	0.596	0.45
16			Bank of Nova Scotia	0.3012	
17	Europe	France	Crédit Agricole	7.56	5.12
18			Societe Generale	2.688	
19	Asia	Israel	Bank Hapoalim	0.5664	0.61
20			Bank Leumi	0.6556	
21	Asia	South Korea	KB Financial Group	0.4484	0.43

22			Shinhan Financial Group	0.4096	
23	Europe	Italy	Unicredit S.P.A	3.752	3.81
24			Intesa Sanpaolo	3.872	
25	Oceania	New Zealand	ANZ Bank New Zealand	0.2852	0.17
26			Kiwi Bank	0.058	
27	South America	Uruguay	Bladex	0.5908	0.86
28			Banco Itaú Uruguay	1.135	

Table 4 Developing

S. NO.	Continent	Country	Bank Name	Weighted Mean NPA Percentage	Mean NPA Percentage
1	Asia	India	State Bank of India	4.604	2.63
2			HDFC	0.6528	
3	Africa	South Africa	Absa Group Limited	2.3068	1.86
4			Firstrand Bank	1.4152	
5	Africa	Nigeria	Zenith Bank	2.2784	2.01
6			Access Bank	1.744	
7	Africa	Zimbabwe	CBZ Bank Zimbabwe	6.6636	5.96
8			ZB Bank Limited Zimbabwe	5.2612	
9	Asia	Malaysia	Maybank Malaysia	1.388	1.48
10			CIMB Bank Malaysia	1.564	
11	Asia	China	Industrial and commercial bank of china	0.9116	0.90
12			China Construction Bank Corporation	0.8868	
13	South America	Chile	Banco Santander Chile	1.3484	1.11
14			Banco de Credito e Inversiones Chile	0.8764	
15	South America	Brazil	Itaú Unibanco Holding	1.532	1.37
16			Brazilian Development Bank	1.2164	
17	South America	Argentina	Banco Patagonia	0.836	1.14
18			Grupo Financiero Galicia	1.4452	
19	Asia	Russia	Sberbank Russia	2.488	3.02
20			VTB Bank Russia	3.556	
21	Africa	Egypt	Commercial International Bank Egypt	3.242	2.33
22			QNB Alahli Bank Egypt	1.4204	
23	Europe	Poland	Powszechna Kasa Oszczędności Bank Poland	3.372	3.41
24			Santander Bank Poland	3.4484	
25	Europe	Croatia	Addiko Bank Croatia	6.56	4.67
26			Erste & Steiermärkische Bank	2.776	
27	North America	Mexico	BBVA Bancomer Mexico	1.36	1.26
28			Banorte Bank Mexico	1.1692	

29	South America	Colombia	Bancolombia S.A.	1.532	1.43
30			Banco de Bogotá Colombia	1.336	

Table 5 Other Parameters (In accordance with Step 6)

	Final Mean NPA Percentage	**Standard Deviation**	**Variance**
Developed	1.2	1.47	2.1609
Developing	2.31	1.45	2.1025

Table 6 T-Value (In accordance with Step 6,7 and 8)

Pooled variance	**Pooled stdev**	**T-VALUE**	**t-value at 95% confidence level (from t-table)**
2.13	0.54	-2.039893696	1.703

5. RESULT AND CONCLUSION

From the calculations, it is observed that the value of $|T| > t$ and in accordance with step 8 the null hypothesis (H_0) taken by us initially would be rejected, thus arriving at the result that the NPAs of developed nations is less than the NPAs of developing nations. This conclusion aims to provide an unprecedented outlook with respect to the development status of a country into the losses faced by their Banks and can serve as a base to the argument that high NPAs of Banks are caused by an indirect effect of low HDI of the country. The causes of NPA mainly depend on some internal and external factors like improper lending practices, banking management, global commodity pricing, etc. With this study, the researchers aimed to bring in a new unconventional factor for the reasons behind NPAs by relating it to the HDI, thereby relating it to the 3 fundamental factors of HDI namely health, education and standard of living.

REFERENCES

[1] ResearchGate Homepage, https://www.researchgate.net/publication/320987078_Essential_Economics/figures?lo=1&utm_source=google&utm_medium=organic, last accessed 2020/05/07.

[2] Investopedia Homepage, https://www.investopedia.com/articles/younginvestors/10/what-is-an-investment.asp, last accessed 2020/05/07.

[3] Our World in Data Homepage, https://ourworldindata.org/human-development-index, last accessed 2020/05/07.

[4] World Bank Homepage, https://data.worldbank.org/country, last accessed 2020/05/07.

[5] Alkire, S. and Deneulin, S. (2009). The human development and capability approach, in (S. Deneulin and L. Shahani, eds.), An Introduction to Human Development and Capability Approach, London: Earthscan.

[6] Ruggeri Laderchi, C., Saithand, R., and Stewart, F. (2003). Does it matter that we do not agree on the definition of poverty? a comparison of four approaches, Oxford Development Studies, 31(3), 243–274.

[7] Ahlem Selma Messai, Fathi Jouini (2013), Micro and Macro Determinants of Non-performing Loans, International Journal of Economics and Financial Issues, 3(4), 852-860.

[8] News18 Homepage, https://www.news18.com/news/india/bad-loans-under-mudra-scheme-on-a-rise-reveals-rti-2441243.html, last accessed 2020/05/07.

[9] Jeevan Nagarkar, Viraja Bhat, Shubham Nagar (2020), Mudra Yojna and Economic Stability in the Indian Context with Special Reference to Gujarat State, International Journal of Innovative Technology and Exploring Engineering, 9(4s), 56-61.

[10] Saibal Ghosh (2005), does leverage influence banks' non-performing loans? Evidence from India, Applied Economics Letter, 12(15), 913-918.

[11] Mandira Sharma, Jesim Pais (2011), Financial Inclusion and Development: A Cross Country Analysis, Journal of International Development, 23(5), 613-766.

[12] Dr. Namita Rajput, Anu Priya Arora, Baljeet Kaur (2012), Management of Non-Performing Assets: A Study of Indian Public Sector Banks, International Journal of Management, IT and Engineering, 2(4), 197-210.

[13] Suman Seth, Antonio Villar (2017), Measuring human development and human deprivations. OPHI Working Papers, 110, 1–30.

[14] Piet H. Renooy (1986), THE INFORMAL ECONOMY; AN URBAN PHENOMENON?, The Netherlands Journal of Housing and Environmental Research, 1(3), 211-225.

[15] Salini R. Chandran, K. Alamelu (2019), NPA and its Impact on Asset Quality- Bankers' Perception, International Journal of Recent Technology and Engineering, 8(2S6), 704-713.

[16] Lalitha, N. (2013), Non-performing Assets Status & Impact: A Comparative Study of Public & Private Sector Banks with Special Reference to SBI, Canara, HDFC and Karur Vysya Banks (Doctoral thesis). Retrieved from http://shodhganga.inflibnet.ac.in/handle/10603/8679.

[17] Raj Kishor Pradhan, Negative Impact of NPA on Indian Economy: An Analysis, In: Proceedings of Public Sector Banks at Crossroads and the Way Forward, Birla Global University, Bhubaneswar (2017).

[18] P.D. Jerome (2007), Regulation of Informal Financial Institutions: A Study of Money Lenders in Kerala, Reserve Bank of India Occasional Papers, 28(1), 1-32.

[19] Tae Kyun Kim (2015), T test as a parametric statistic, Korean Journal of Anesthesiology, 68(6), 540-546.

[20] Wikipedia Homepage, https://en.wikipedia.org/wiki/Student%27s_t-test#History, last accessed 2020/05/07.

[21] Ravedata Homepage, https://ravedata.in/statistics/t-test/, last accessed 2020/05/07.

[22] T table Homepage, http://www.ttable.org/, last accessed 2020/05/07.

View publication stats



EasyChair Preprint

№ 4109

Study of Model of Crankshaft and Tools

Siddharth Boon, R. C. Singh and Rajiv Chaudhary

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

August 30, 2020

Study of Model of Crankshaft and Tools

[1]SIDDHARTH BOON,

Scholar, Department of Mechanical Engineering, Delhi Technological University

[2]Dr. R.C SINGH, [3]Dr. Rajiv Chaudhary

Profesor, Department of Mechanical Engineering, Delhi Technological University

Abstract:

Crankshaft is enormous volume creation segment with a mind-boggling geometry in the Internal Combustion (I.C) Engine. This proselytes the responding relocation of the cylinder in to a rotational movement of the wrench. An endeavor is made in this paper to examine the Static investigation on a crankshaft from a solitary chamber 4-stroke I.C Engine. The displaying of the crankshaft is made utilizing CATIA-V5 Software. Limited component examination is performed to acquire the variety of stress at basic areas of the driving rod utilizing the ANSYS programming and applying the limit conditions. At that point the outcomes are drawn Von-misses stress prompted in the crankshaft is 15.83Mpa and shear stress is actuated in the crankshaft is 8.271Mpa. The Theoretical outcomes are acquired von-misses stress is 19.6Mpa, shear stress is 9.28Mpa. The approval of model is contrasted and the Theoretical and FEA aftereffects of Von-misses stress and shear stress are inside the cutoff points. Further it very well may be stretched out for the various materials and dynamic examination, advancement of driving rod.

Keywords: Crankshaft; Engine, Stroke, ANSYS, CAD

Introduction:

A crankshaft is a rotating shaft which (in conjunction with the connecting rods) converts reciprocating motion of the pistons into rotational motion. Crankshafts are commonly used in internal combustion engines and consist of a series of cranks and crankpins to which the connecting rods are attached.

The crankshaft rotates within the engine block through use of main bearings, and the crankpins rotate within the connecting rods using rod bearings. Crankshafts are usually made from metal, with most modern crankshafts being constructed using forged steel.

The concept of using crankshaft is to change these sudden displacements to as smooth rotary output, which is the input to many devices such as generators, pumps and compressors. It

should also be stated that the use of a flywheel helps in smoothing the shocks. Crankshaft experiences large forces from gas combustion. This force is applied to the top of the piston and since the connecting rod connects the piston to the crank shaft, the force will be transmitted to the crankshaft. The magnitude of the forces depends on many factors which consist of crank radius, connecting rod dimensions, and weight of the connecting rod, piston, piston rings, and pin. Combustion and inertia forces acting on the crankshaft

- Torsional Load
- Bending Load

Crankshaft must be strong enough to take the downward force of the power stroke without excessive bending so the reliability and life of the internal combustion engine depend on the strength of the crankshaft largely. The crank pin is like a built in beam with a distributed load along its length that varies with crank positions. Each web is like a cantilever beam subjected to bending and twisting.

Cylinder Arrangement

The crankshaft is found beneath the chambers of a vehicle's engine. On V-type engines it is found at the base yet on level engines it is found between the chamber banks. Engine vehicles may have 3 to 12 chambers inside the engine albeit most have four. Inside every chamber is a cylinder which goes here and there the chamber. All the engine's cylinders are associated with the crankshaft by singular poles. The chambers work in show just as with other engine parts. This is alluded to as the four-stroke cycle and happens in every one of the four chambers. This cycle is the thing that drives a vehicle's engine.

Four-Stroke Cycle

The four strokes allude to allow, pressure, force and fumes. On the admission stroke, the cylinder begins down as the admission valve opens to permit air and fuel into the chamber. When the cylinder shows up at the base of the admission stroke, it triggers conclusion of the admission valve. The air-fuel blend is held in the chamber. This blend is compacted seriously by the cylinder as it climbs. The chamber substance is touched off by the flash fitting during which process extension happens. The ignition procedure brings down the cylinder which turns the crank to yield capacity to drive the vehicle. The fumes valve at that point opens to discharge the fumes once the cylinder gets to the base of the chamber.

Crankshaft-Camshaft Operations

The crank moves the cylinders all over inside the chambers. The development of the cylinders is directed by the crank. A part known as the camshaft additionally guarantees that the cylinders work appropriately. At whatever point the crank pivots, the camshaft should likewise turn alongside it. This is on the grounds that the two segments are connected together. The two engine parts have a synchronized development. At the point when the camshaft pivots it makes the admission and outtake valves open. This permits a progression of air which is imperative to cause blasts in the chamber. Blasts are made inside the chambers in the engine. The blasts apply pressure on the cylinders with the goal that they keep up their development. These blasts bring about development of the wheels. The moving cylinders offer ascent to jerky developments. The flywheel which is found toward the finish of the shaft assists with facilitating the sporadic development. At the point when the shaft moves, it makes the flywheel embrace a round movement. Scores in the flywheel help it to accomplish a progressively ordinary movement. This movement in the long run makes the vehicle's wheels turn since the flywheel is associated with other engine parts.

Working:

Force from the consumed gases in the ignition chamber is conveyed to the crankshaft through the cylinder, cylinder pin and associating bar. The crankshaft changes responding movement of the cylinder in chamber to the turning movement of the flywheel. Change of movement is executed by utilization of the counterbalance in the crankshaft. Each balance some portion of the crankshaft has a direction surface known as a crank pin to which the interfacing bar is joined. Crank-through is the counterbalanced from the crankshaft place line. The stroke of the cylinder is constrained by the toss of the crankshaft. The burning power is moved to the crank-toss after the crankshaft has moved past top right on target to deliver turning exertion or force, which pivots the crankshaft. In this manner all the engine power is conveyed through the crankshaft. The cam-shaft is turned by the crankshaft through apparatuses utilizing chain driven or belt driven sprockets. The cam-shaft drive is coordinated for opening of the valves corresponding to the cylinder position. The crankshaft turns in primary heading, which are part down the middle for get together around the crankshaft fundamental bearing diaries.

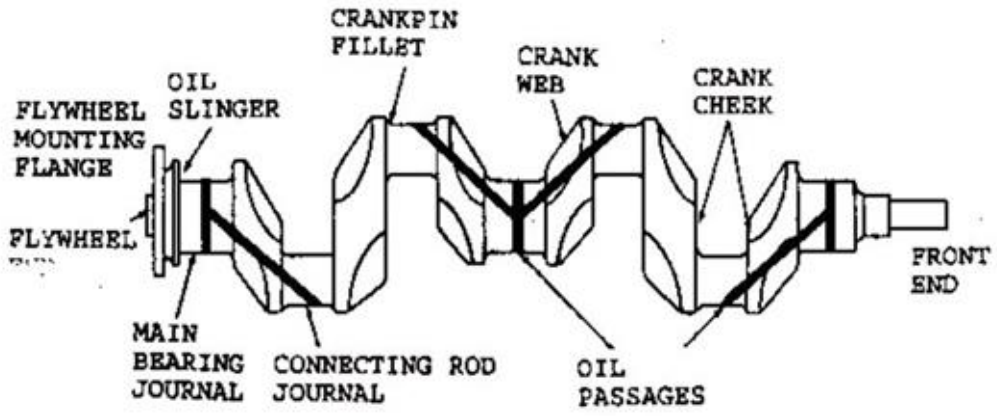


Figure -1

Applications of CAD:

Computer Aided Design (CAD) is a sort of computer-based apparatus utilized for drafting and designing. Computer aided design is helpful in different designing fields, for example, engineering, mechanical and electrical fields being some of them. This is a sort of programming, which empowers clients to make quick and exact drawings and unpleasant sketch plans of primary items. It gives an adaptable example in the drawing procedure that clients can modify as indicated by their necessary measurements with negligible endeavors.

Computer aided design isn't just made for specialists explicitly yet has the assorted variety to engage a wide range of designing fans. This product has all worked in highlights according to clients need and accompanies numerous layouts and images, for designing and drafting purposes, which gives it a wide zone of utilization. It is the essential geometry-composing instrument utilized for all 2D and 3D designing purposes.

Application of CATIA

CATIA is an abbreviation for Computer Aided Three-dimensional Interactive Application. It is one of the main 3D programming utilized by associations in various businesses extending from aviation, vehicle to customer items.

CATIA gives the ability to picture designs in 3D. At the point when it was presented, this idea was imaginative. Since Dassault Systems didn't have a skill in advertising, they had income sharing tie-up with IBM which demonstrated incredibly productive to both the organizations to showcase CATIA. In the beginning phases, CATIA was broadly utilized in the design of the Mirage airplanes; anyway the capability of the product before long settled on it a mainstream decision in the car division also. As CATIA was acknowledged by

increasingly fabricating organizations, Dassault changed the item grouping from CAD/CAM programming to Project Lifecycle Management. The organization likewise extended the extent of the product.

CATIA can be utilized at various phases of the design - ideate, draw, test and repeat. The product accompanies various workbenches ("modules") that permit CATIA to be utilized across changed businesses – from parts design, surface design and get together to sheet metal design. CATIA can likewise be utilized for CNC.

Simulation:

Ansys Simulation is a design investigation framework completely coordinated with Solid works. Strong works Simulation gives reenactment answers for direct and nonlinear static, recurrence, clasping, warm, weakness, pressure vessel, drop test, straight and nonlinear dynamic and streamlining examinations.



Figure 2: simulation example

Structural Analysis

Table 1: Properties of material

Material	Density(g/cm^3)	Youngs modulus(Gpa)	Poissions ratio
Aluminum alloy	2.6898	68.3	0.34
Titanium alloy	4.62	96	0.36
Magnesium alloy	1.8	45	0.35

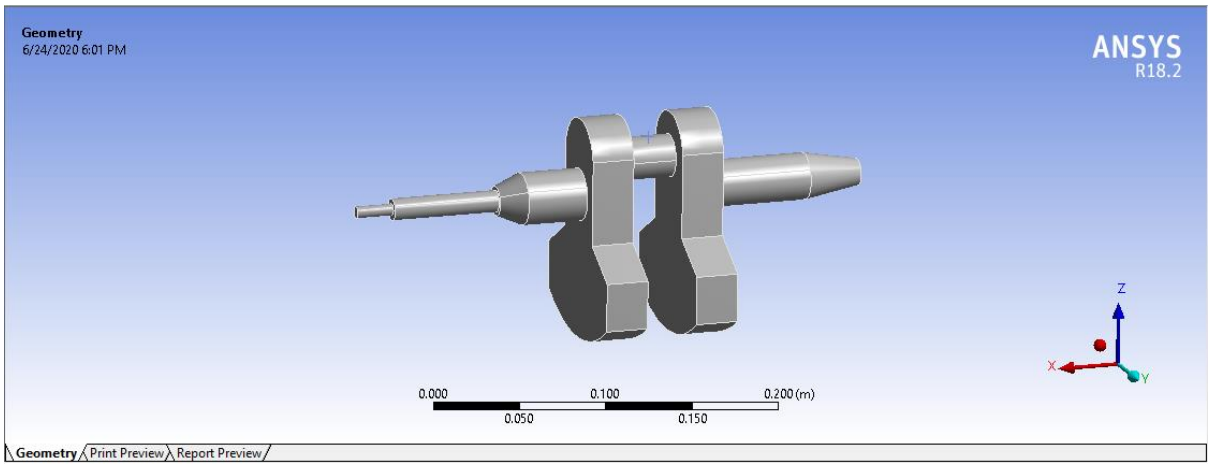


Figure 3: Model

Results:

Table 2: results for all analysis

Material	Total deformation (M)	Strain	Stress (pascals)
Aluminum alloy	2.0643e-7	4.6718e-6	3.2003e5
Titanium alloy	1.5034e-7	3.3248e-6	3.1719e5
Magnesium alloy	3.2248e-7	7.1889e-6	3.1815e5

1. Modeling of crank shaft is done in catia v5 design programming by utilizing different orders
2. From the tables 2 it is presumed that the titanium amalgam is demonstrating proficient outcomes
3. Hence titanium composite is best among the three applied materials
4. Maximum stress, twisting and most extreme strain are noted and arranged

Conclusion:

In this research paper we studied the model of the crankshaft and the brief introduction of the software used to design the model of the crankshaft, we also studied the simulation tool "ANSYS" so that we have simulated the various factors like Static Analysis, Thermal Stress Analysis, Frequency examination, Dynamic investigation and Bulking investigation etc., so that we can go for the best. Modeling of crank shaft can be done in "Catia v5" design programming by utilizing different orders.

REFERENCES

[1.] Rincle Garg, Sunil Baghla, "Finite element analysis and optimization of crankshaft", International Journal of Engineering and Management Reaserch, vol-2,Issue-6,ISSN: 2250-0758, Pages:26-31, December 2012.

[2.] C.M Balamurugan, R. Krishnaraj, Dr.M.sakhivel, K.kanthavel, Deepan Marudachalam M.G, R.Palani, "Computer Aided modelling and optimization of Crankshaft", International Journal of scientific and Engineering Reaserach, Vol-2, issue-8, ISSN:2229- 5518, August-2011.

[3.] Gu Yingkui, Zhou Zhibo, "Strength Analysis of Diesel Engine Crankshaft Based on PRO/E and ANSYS", Third International Conference on Measuring Technology and Mechatronics Automation, 2011.

[4.] Abhishek choubey, Jamin Brahmbhatt, "Design and Analysis of Crankshaft for single cylinder 4- stroke engine", International Journal of Advanced Engineering Reaserch and studies, vol-1, issue-4, ISSN:2249-8974, pages: 88-90, July-sept 2012.

[5.] R.J Deshbhratar, Y.R Suple, " Analysis and optimization of Crankshaft using FEM", International Journal of Modern Engineering Reasearch, vol-2, issue-5, ISSN:2249-6645, pages:3086-3088, Sept-Oct 2012.

[6.] Farzin H. Montazersadgh and Ali Fatemi " Stress Analysis and Optimization of Crankshafts Subjected to Dynamic Loading", AISI, August 2007.

[7]Xiaorong Zhou., Ganwei Cai., Zhuan Zhang. Zhongqing Cheng., 2009, "Analysis on Dynamic Characteristics of Internal Combustion Engine INTERNATIONAL JOURNAL OF PROFESSIONAL ENGINEERING STUDIES Volume VIII /Issue 5 / AUG 2017 IJPRES Crankshaft System," International Conference on Measuring Technology and Mechatronics Automation.

[8]. Farzin H. Montazersadgh and Ali Fatemi., 2007,"Dynamic Load and Stress Analysis of a Crankshaft," SAE Technical Paper No. 010258, Society of Automotive Engineers

J Mater Sci: Mater Electron

Tailoring of electrical and optical properties of regenerated silk fibroin films with metal oxides

Reetu Yadav[1] and Roli Purwar[1,*]

[1] *Discipline of Polymer Science and Chemical Technology, Department of Applied Chemistry, Delhi Technological University, Shahbad Daulatpur, Bawana Road, Delhi 110042, India*

Received: 2 July 2020
Accepted: 23 August 2020

© Springer Science+Business Media, LLC, part of Springer Nature 2020

ABSTRACT

In this study, silk fibroin was used as a matrix to incorporate the metal oxide nanoparticles i.e. zinc oxide (ZnO) and copper oxide (CuO) to form transparent nanocomposite films by solvent casting method. The effect of metal oxide nanoparticles and their concentrations on the conductivity of silk fibroin film were detected by a four-point probe instrument. As the thickness of the films was reduced, the conductivity increased. Hall Effect experiment was carried out to determine the carrier (electron or holes) concentration and identification of the semiconducting behaviour of nanocomposite films. Pure silk fibroin (SF) film showed p-type semiconducting behaviour. Incorporation of zinc oxide and copper oxide nanoparticles in the silk fibroin matrix makes it n-type and p-type semiconductor, respectively. Uniform dispersion of metal oxide nanoparticles in the silk fibroin matrix was observed through field emission scanning electron microscopy. The UV–Vis spectra of nanocomposite films showed 82–84% transmittance and reduction in refractive index. Addition of metal oxide nanoparticles reduced the β sheet content and percentage crystallinity of silk fibroin as observed by FTIR and XRD, respectively. The structural alterations at the molecular level reveal their relationship with electrical and optical properties.

1 Introduction

The aspiration to fabricate electronic devices using green materials has opened new avenues in the field of biomaterial research [1, 2]. *Bombyx mori*-derived silk fibroin, one of the most abundant protein, has recently been found to be a promising choice for organic electronics due to its solution processability, tailored mechanical properties, thermal stability, optical properties, biodegradability, and biocompatibility [3, 4]. Regenerated silk fibroin films are reported to be highly insulating with electrical conductivity in the range of 10^{-9} to 10^{-15} S cm^{-1} [1, 5]. Recently, efforts have been given by various research groups to enhance the conductivity of the regenerated silk fibroin films to develop material for biosensors, flexible organic light-emitting diodes, and wearable electronic textiles [6–11]. The conductivity

Address correspondence to E-mail: roli.purwar@dtu.ac.in

https://doi.org/10.1007/s10854-020-04332-4
Published online: 08 September 2020

of the regenerated silk fibroin films can be enhanced either by blending with conducting polymers such as polypyrrole, polyaniline, or by the formation of nanocomposites using conductive nanostructure materials such as single-wall carbon nanotubes, multiwalled carbon nanotubes, graphene, silver and gold nanoparticles, silver nanowires, and hydrotalcite [7–10, 12–17]. The addition of these nanomaterial structures has enhanced the conductivity of regenerated silk fibroin films in the range of 10^{-3} to 10^{-9} S cm^{-1} i.e. semiconductor range. The regenerated silk fibroin film cast over patterned silver nanowire showed 80% optical transparency with good conductivity [18]. The formation of black or grey colour limits the applications of silk fibroin/carbon-based nanostructure conductive nanocomposite films where both electrical and optical properties are required. Further, there is a need to develop semiconducting regenerated silk fibroin films (both n-type and p-type) with optical transparency to be used in the field of optoelectronics.

Metal oxide nanoparticles such as zinc oxide (ZnO) and copper oxide (CuO) have outstanding electrical, magnetic, thermal, and optical properties. These metal oxide nanoparticles have been used to tune the conducting properties of polymer nanocomposites [19–21]. Zinc oxide is an n-type semiconductor material with a wide bandgap (3.37 eV) and a large excitonic binding energy (5.3 eV) have diverse properties such as piezoelectricity, chemical stability, biocompatibility, optical transparency in the visible region, and high voltage–current non-linearity [22]. Zinc oxide has been used with various polymers like epoxy, PMMA, polycarbonate, starch, LDPE for diode, and memory cell applications [20, 23–25]. Copper oxide is a monoclinic p-type semiconductor with narrow bandgap energy of 1.5–1.8 eV and 6 eV excitation binding energy, having good mechanical flexibility and environmental stability. Its resistivity could be controlled with acid–base doping [26]. Polymers like polyvinyl alcohol, chitosan, and polyaniline have been used with copper oxide to form nanocomposites for nanofluids, lightweight battery electrodes, electromagnetic shielding devices, anticorrosion coatings, solar cells, photodetectors, and sensors [21, 27, 28].

Since different metal oxides have different semiconducting behaviour (n-type or p-type), it is interesting to examine how these metal oxides affect the electrical, optical, and structural properties of regenerated silk fibroin films. In this study, two different types of metal oxides: Zinc oxide (n-type) and Copper oxide (p-type) have been chosen to study their effect on regenerated silk fibroin films. Different concentrations of metal oxides are used to fabricate regenerate silk fibroin/metal oxides nanocomposite films using a solution casting method with varied thickness. The concentration of metal oxide nanoparticles and thickness of the films are optimized based on the electrical conductivity. The optimized silk fibroin/metal oxides nanocomposite films are evaluated for the type of semiconducting behaviour, optical, and structural properties.

2 Experimental

2.1 Materials

Mulberry (*Bombyx mori*) silk cocoons were supplied from Central Sericulture Research and Training Institute, Mysore, India. Calcium chloride fused LR was purchased from Thomas Baker (Chemicals) Pvt. Ltd. Glycerol (99.5% purity) was supplied from Thermos Fisher Scientific India Pvt. Ltd. Methanol (AR) and formic acid (AR) 98% purity were obtained from Avantor performance materials India Ltd. Zinc oxide nanopowder (Type I) (Average particle size: 30 nm) and Copper oxide nanopowder (Average particle size: 40 nm) were purchased from Sisco Research Laboratories Pvt. Ltd. Milli-Q water was used in the coagulation process.

2.2 Fabrication of silk fibroin/metal oxide nanocomposite films

2.2.1 *Extraction of silk fibroin (SF)*

Degumming of silk cocoons was carried out according to the previously published procedures [29]. In brief, the *Bombyx mori* cocoons were cut into small pieces, vacuum dried, and placed in boiling solution of Na_2CO_3 (0.02 M) (1 L) for 2–3 h to remove the trace of sericin or other impurities completely. The obtained silk fibroins were rinsed 3–4 times in hot water (2 L) for 30 min and dried overnight in an oven at 70 °C.

2.2.2 *Fabrication of films*

Metal oxide nanoparticles i.e. zinc oxide and copper oxide were sonicated in 2% $CaCl_2$/formic acid (w/v) solution for 30 min. Sonicated solution was used to dissolve the silk fibroin at room temperature by continuous stirring at 500 rpm for 3–4 h. The concentration of silk fibroin was kept 10 wt% (w/v) in the formic acid solution. Glycerol (15 wt% w/v) was added in the silk fibroin solution to induce flexibility in the films [30]. Concentrations of metal oxide nanoparticles in the silk fibroin solution were varying as 0.0, 0.1, 0.25, 0.5, 0.75, and 1.0% w/w. Films were prepared by the solvent casting method. A variable amount of solution was poured on Petri plates in order to formulate films of desired approximate thicknesses i.e. 30 µm and 70 µm. Films were dried overnight then dried films immersed in 100% methanol for 4 h followed by 80% methanol (24 h) for coagulation. The coagulated films were then kept for water annealing (15 min). Then the films were stored in refrigerator at 4 °C. Figure 1 shows the schematic diagram for the preparation of transparent silk fibroin/metal oxides nanocomposite films. The films were abbreviated as Pure SF, SF/ZnO, and SF/CuO for regenerated pure silk fibroin, zinc oxide, and copper oxide nanoparticles-incorporated silk fibroin films, respectively. The thicknesses of pure SF and SF/ZnO and SF/CuO nanocomposite films were measured by using the digital thickness gauge micrometer of MGW Precision OMC25 (LC 0.001 mm).

2.3 Electrical and optical properties of silk fibroin/metal oxides nanocomposite films

Conductivity measurements of the pure SF, SF/ZnO, and SF/CuO nanocomposite films were measured by using the four-point probe method of SES Instruments. Sheet resistivity of the samples was recorded by sample size (2 × 2) cm^2 placed on the four-point probe in which current was passed through the outer probe and voltage was measured through the inner probes.

The sheet resistivity and conductivity of the samples were calculated based on current–voltage relationships using the following formulas [31]

$$\rho = \frac{\pi}{\ln(2)}\left(\frac{V}{I}\right)t = 4.532(R)t \tag{1}$$

$$\sigma = \frac{1}{\rho}, \tag{2}$$

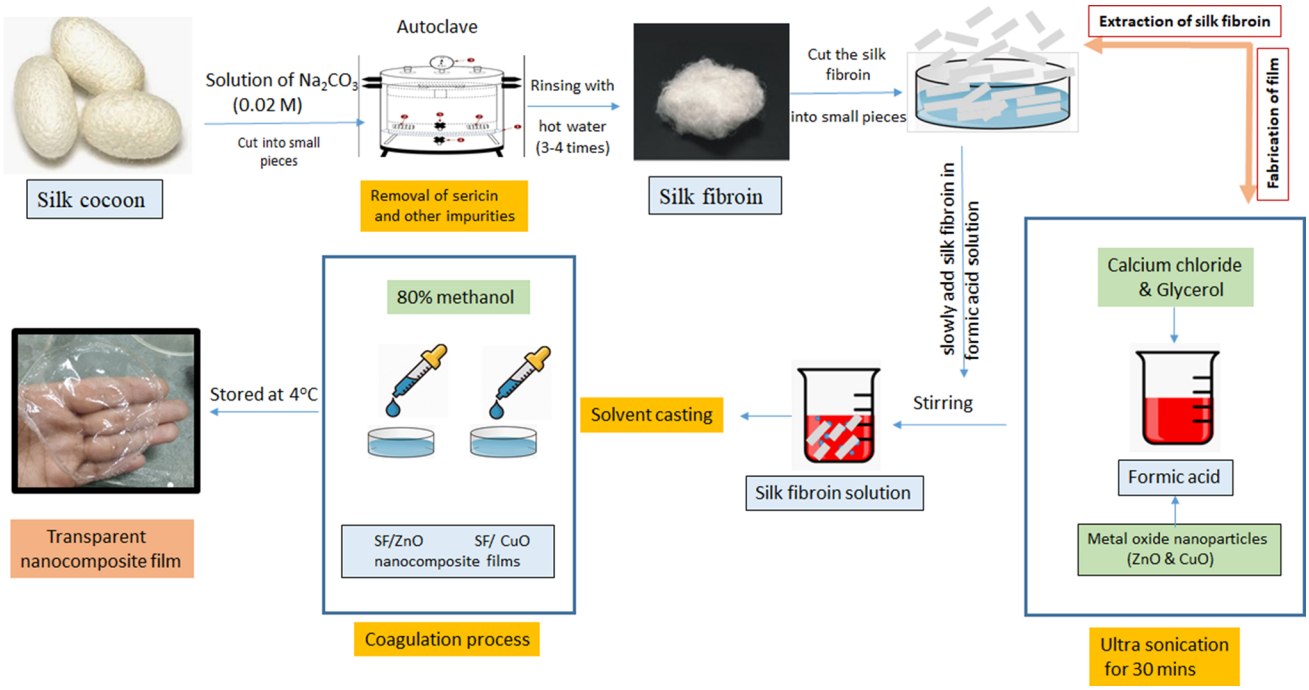


Fig. 1 Extraction of silk fibroin and fabrication process of transparent silk fibroin/metal oxide nanocomposite films

where R is the resistance (ohm), V is the voltage (V), I is the current (A), t is the thickness in cm, ρ is the sheet resistivity (Ω/square), σ is the conductivity (S cm^{-1})

The semiconducting material characterization of pure SF and SF/ZnO and SF/CuO nanocomposite films was carried out using the Hall effect measurement instrument. To determine the hall voltage, hall coefficient, and carrier concentration, film samples were analysed at the dry condition at room temperature by using SES Grey Magnetoresistance, MRX-01, Roorkee, India. The hall voltage value was obtained through the Hall Effect experiment where probe current was around 2 A and magnetic field strength was kept at 0.04 T. Hall coefficient and carrier concentration were determined by following formulas [32]

$$n = \frac{IB}{qv_{\mathrm{H}}d} \quad R_{\mathrm{H}} = -\frac{1}{nq}, \tag{3}$$

where V_{H} is the hall voltage (mV), R_{H} is the hall coefficient (m^3/C), I is the current flowing through the material (mA) obtained through experiment, B is the magnetic field strength (T), q is the charge (C), n is the carrier concentration, i.e. number of mobile charge carriers per unit volume (cm^3), d is the thickness of the material in cm.

The optical properties of pure SF and SF/ZnO and SF/CuO nanocomposite films were analysed by using the UV–Vis spectrophotometer of carry 300 Bio spectrophotometer (Agilent Technologies, Inc. U.S). In the wavelength range of 200–800 nm, the structural modifications induced by metal oxide nanoparticles can cause a change in the molecular structure of the silk fibroin and hence a modification within the optical bandgap was expected. The information about the optical bandgap was obtained from the absorption edge of the UV–Vis spectra of the silk fibroin. The optical absorption coefficient (α) of pure SF and SF/ZnO and SF/CuO nanocomposite films was calculated from the optical absorbance (A). To determine the optical bandgap energy (E_{g}) and Urbach energy (E_{u}) of pure SF and SF/ZnO and SF/CuO nanocomposite films, the following equations were used [33]

$$\text{Absorption coefficient,} \quad \alpha = 2.303\frac{A}{(X)hv} \tag{4}$$

$$\text{Optical bandgap energy,} \quad \alpha(hv) = B\frac{(hv - E_{\mathrm{g}})^r}{hv} \tag{5}$$

$$\text{Urbach energy,} \quad \alpha(hv) = \alpha_0 \exp(hv - E_{\mathrm{u}}), \tag{6}$$

where α is absorption coefficient (cm^{-1}), (hv) is the energy of the incident photons (in eV), A is absorbance defined as $A = \log(I_0/I)$, where I_0 and I are the intensity of the incident and transmitted beam, respectively, X is the thickness of the film (cm), E_{g} is the optical bandgap energy which is obtained by variation of $(\alpha hv)^2$ as a function of photon energy (hv) and determined by extrapolation of the linear region of the spectra to intersect at $(\alpha hv)^2 = 0$, r is an empirical exponent which characterizes the type of optical transition during the absorption process. The exponent r has the value of 1/2 for allowed direct transition, B depends on the transition probability and can be assumed to be constant within the optical frequency range, α_0 is a constant, E_{u} is an Urbach energy or the band tail energy calculated from the slope of the line on plotting ln (α) against the incident photon energy (hv). Values were obtained by taking the reciprocal of the line slopes, v is the frequency of radiation, and h is Planck's constant.

The refractive index of polymer films can be determined by using UV–Vis spectrophotometer by measuring the absorption coefficient (α) and extinction coefficient (k) at a particular wavelength. The refractive index (n) values for pure SF and SF/ZnO and SF/CuO nanocomposite films were determined at 550 nm wavelength using the following equation [33]:

$$\text{Refractive index,} \quad n = \left[\frac{4r}{(r-1)^2} - k^2\right]^{0.5} - \frac{r+1}{r-1}, \tag{7}$$

where r has a value of 1/2 for allowed direct transition and k is the extinction coefficient which is a measure of the fraction of light lost due to scattering and absorption per unit distance of the penetration medium. It can be determined by using the relation [33]:

$$k = \frac{\alpha\lambda}{4\pi}, \tag{8}$$

where k is the extinction coefficient, α is the absorption coefficient, and λ is the wavelength (nm).

2.4 Characterization of silk fibroin/metal oxides nanocomposite films

Surface morphology of the fabricated pure SF and SF/ZnO and SF/CuO nanocomposite films were characterized by using the field emission scanning electron microscope (FESEM) in high vacuum accelerated voltage of 15 kV. FESEM images were acquired by JEOL Model JSM-7900F prime with oxford EDS system, Netherland. The vacuum-dried samples were sputter-coated with platinum before mounting on sample holder of FESEM-EDS instrument.

The absorption spectra of pure SF and SF/ZnO and SF/CuO nanocomposite films were obtained through Fourier transform infrared (FTIR) spectroscopy using the ATR plate. Spectra were recorded in the range of 400–4000 cm^{-1} by using Bruker, Tensor 37 model. To elucidate the secondary structure and β-sheet content in the silk fibroin films, deconvolution method was used on the absorption spectra. Deconvolution of amide I band was carried out using Peak Fit. The numbers and positions of peaks were defined from the second-derivative results of the spectra. Peaks were fixed during the deconvolution process and data minimized by the Chi-squared criterion. A Gaussian model with integral type was applied for the band shape and bandwidth to adjust by the origin pro 9.0 software [34].

Wide-angle X-ray diffraction was used to calculate the chemical composition and crystallinity of pure SF and SF/ZnO and SF/CuO nanocomposite films. XRD spectra were obtained from 10° to 40° angle by PANalytical, X'Pert PRO model, Netherlands. To determine the β-sheet crystallinity in the silk fibroin films, the deconvolution method was used on the XRD spectra. Deconvolution of spectra was carried out using Peak Fit. Data were minimized by the chi-squared criterion. A Gaussian model with integral type was applied for the band shape and bandwidth to adjust by the origin pro 9.0 software [34].

3 Results and discussion

3.1 Electrical conductivity of silk fibroin/metal oxides nanocomposite films

Zinc oxide is a stable, non-toxic, environment-friendly metal oxide which also possesses a wide bandgap and low work function (3.37 eV). The n-type nature of this metal oxide is due to the oxygen vacancies and interstitial atoms which induce shallow electron donor states that contribute to the conductivity. Copper oxide has p-type semiconductor behaviour with a narrow bandgap of 1.5 eV and high work function (5.3 eV) [22]. The effect of concentration of metal oxide nanoparticles on the conductivity of the silk fibroin/metal oxide nanocomposite films was measured at different thicknesses with a four-probe instrument and the results are shown in Fig. 2. Conductivity of pure silk fibroin (pure SF) films with 30 μm and 70 μm thickness are 35.86×10^{-6} S cm^{-1} and 32.65×10^{-6} S cm^{-1}, respectively. The conductivity of regenerated silk fibroin (pure SF) film is reported in the range of 10^{-9} S cm^{-1} to 10^{-15} S cm^{-1} [1, 5]. The electrical conductivity of silk fibroin film depends upon the ionic impurities present during the preparation of the film, β sheet content, and method

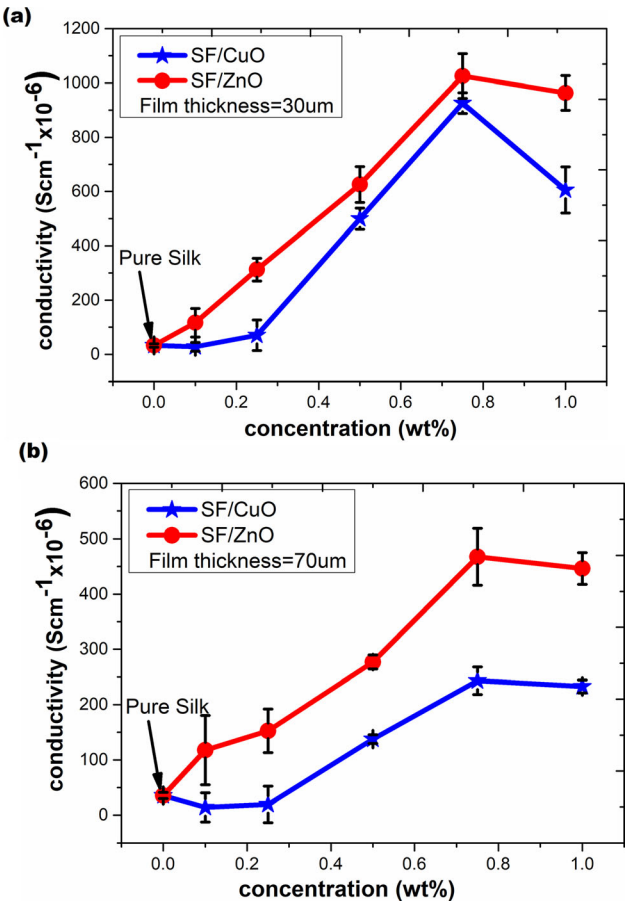


Fig. 2 Electrical conductivity of pure silk film and SF/ZnO and SF/CuO nanocomposite films **a** conductivity of film thickness 30 μm, **b** conductivity of film thickness 70 μm

of measurement [6]. In this case, ionic impurities such as calcium chloride and glycerol were added during the formation of the film. Therefore, pure SF films also show higher conductivity as compared to values reported in the literature. As the concentration of metal oxide nanoparticles increases, the conductivity of nanocomposite films increases up to 0.75% w/w after that there is no improvement observed (Fig. 2a, b). The nanocomposite films also show higher conductivity at lower thickness. Thickness of the films has an enormous impact on electrical conductivity. Similar results have been reported by Liang et al., for the conduction current of ultra-thin films of polyimides due to better orientation of polymer chain at lower film thickness [35]. The addition of metal oxide nanoparticles enhanced the conductivity of nanocomposite films from 10^{-6} to 10^{-3} S cm^{-1} with 30 μm thickness. It is also observed that SF/ZnO nanocomposite films have higher conductivity as compared to that of SF/CuO nanocomposite films (Fig. 2a, b). The addition of metal nanoparticles in the silk fibroin matrix may affect the carrier (electron or hole) mobility in the matrix and change the conductivity of silk fibroin/metal oxides nanocomposite films. Dispersion of metal oxide nanoparticles, the type of carrier concentration, and β sheet content are some of the factors which may also play a major role for difference in the conductivity of two different metal oxides-incorporated nanocomposite silk fibroin films.

Silk fibroin/metal oxides films with 30 μm thickness having 0.75% w/w metal oxides concentration have shown the best electrical conductivity therefore, further evaluations are carried out on these films. The nanocomposite films are abbreviated as SF/ZnO (0.75) and SF/CuO (0.75).

To understand the semiconducting behaviour and carrier concentration present in nanocomposite films, the Hall Effect experiment was carried out. The Hall Effect measurements are important to semiconductor material characterization like hall voltage, hall coefficient, conductivity type, and carrier concentration. Hall Effect is the production of a voltage difference (the Hall voltage) across an electrical conductor or semiconductor, transverse to an electric current in the conductor, and an applied magnetic field perpendicular to the current. The Hall voltage is directly proportional to the current flowing through the material, and the magnetic field strength, and it is inversely proportional to the number of mobile charges in the material, and the thickness of the material. So, in order to produce a large Hall voltage, we need to use a thin material with few mobile charges per unit volume. This experiment explained if the Hall voltage produced is positive then the material is said to be p-type and if the Hall voltage produced negative then the material is said to be n-type [32, 36]. The results of the Hall Effect experiment are shown in Table 1. It is observed that pure SF film produced the positive Hall voltage ($+ 26.5 \pm 0.4$) with Hall coefficient of $- 96.1 \pm 1.8$ m^3/C which depicts the p-type semiconducting behaviour. On incorporating the p-type metal oxide i.e. CuO, SF/CuO (0.75) film showed the positive hall voltage ($+ 14.1 \pm 0.6$ mV) with enhancement in Hall coefficient value of $- 48.3 \pm 0.8$ m^3/C. It is depicted from the observation that holes are the charge carriers in pure SF and SF/CuO (0.75). On incorporation of n-type metal oxide nanoparticles i.e. ZnO, the SF/ZnO (0.75) nanocomposite film showed the negative Hall voltage of $- 0.3 \pm 0.1$ mV. It suggests that the n-type behaviour of the SF/ZnO (0.75) nanocomposite films and the charge carriers are electrons. The carrier concentrations are also found highest in the SF/ZnO (0.75) nanocomposite films. The values of carrier concentration are found to be in accordance with the electrical conductivity behaviour of nanocomposite and pure silk fibroin films. The movement of the charge carriers (holes or electrons) in the applied electrical field is dependent on the interaction between the silk fibroin and metal oxide nanoparticles. Sangappa et al. have reported the conductivity in silk fibroin/gold nanocomposite films with charge hopping mechanism. The hydroxyl groups of silk fibroin and gold nanoparticles are forming charge transfer complex which reduces the barrier height between the tapping sites and provides conducting pathways through the amorphous region of silk fibroin. The protein conformation plays a critical role in conducting pathways in the silk fibroin matrix [5]. Shivananda et al. have described that the increase in electrical conductivity of silk fibroin on the addition of silver nanoparticles is due to the accumulation of mobile charge at the interface which forms large dipoles on silver metal particles [8]. There is a possibility that a similar charge transfer complex was formed in nanocomposite films on incorporating metal oxide nanoparticles in silk fibroin which is highly influenced by the alteration of protein

Table 1 Semiconducting behaviour of pure SF and SF/ZnO and SF/CuO nanocomposite films in terms of Hall coefficient and carrier concentration

Sample	Hall voltage (mV)	Hall coefficient (m^3/C)	Carrier concentration (Cm^3)	Type of semiconducting behaviour
Pure SF	(+) 26.5 ± 0.4	(−)96.1 ± 1.8	$6.4 \times 10^{16} \pm 1.2$	p- type
SF/ZnO (0.75)	(−) 0.3 ± 0.1	(+) 1.0 ± 0.4	$573.6 \times 10^{16} \pm 4.7$	n-type
SF/CuO (0.75)	(+) 14.1 ± 0.6	(−) 48.3 ± 0.8	$12.9 \times 10^{16} \pm 3.9$	p-type

conformation in nanocomposite films that facilitates the charge hopping mechanism.

3.2 Optical properties of silk fibroin/metal oxides nanocomposite films

In the ultraviolet region, when the light energy is absorbed by the polymeric material, the transition of electrons takes place from the ground state to excited energy states at different orbitals. The electronic transitions are involved $\sigma \rightarrow \sigma^*$, $n \rightarrow \sigma^*$, $n \rightarrow \pi^*$, and $\pi \rightarrow \pi^*$ [33]. The UV–Vis absorbance spectra of pure SF and SF/ZnO and SF/CuO nanocomposite films are presented in Fig. 3a. The recorded spectra of pure SF film display absorbance at 300 nm which is assigned to $\pi \rightarrow \pi^*$ transition of the Tyr, Phe, and Try residues in the SF molecular chain [33]. Transmittance spectra of these films are also recorded and presented in Fig. 3b. Pure SF films showed 74% transmittance at 550 nm while the SF/ZnO (0.75) and SF/CuO (0.75) nanocomposite films showed 84% and 82% transmittance, respectively. This increase in optical transmittance may be attributed to the occurrence of structural changes on the incorporation of metal oxide nanoparticles in silk fibroin. Bucciarelli et al. have reported that high β sheet content in silk fibroin leads to lower transmittance due to scattering [37]. It can be hypothesized that metal oxide nanoparticles are bringing some change in protein conformation which caused the enhancement in optical transmittance of nanocomposite films.

The information about the optical bandgap is obtained from the absorption edge of the UV–Vis spectra of the pure SF and nanocomposites samples. The optical absorption coefficient (α) of films was calculated from the optical absorbance A, using Eq. 4 (given in the materials and method section). The optical bandgap energy was obtained by plotting the variation of $(\alpha h\nu)^2$ as a function of photon energy ($h\nu$) and determined by extrapolation of the linear region

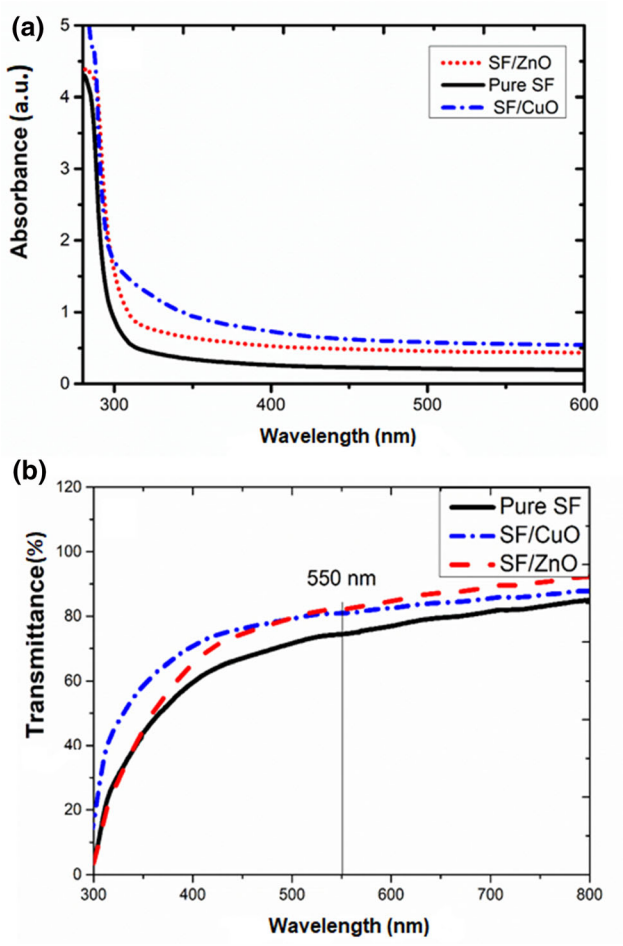


Fig. 3 UV–Vis (**a**) absorbance spectra (**b**) transmittance spectra of pure SF film and SF/ZnO (0.75) and SF/CuO (0.75) nanocomposite films with a thickness of 30 μm

of the spectra to intersect at $(\alpha h\nu)^2 = 0$ (Fig. 4a). The pure SF film of 30 μm thickness has an optical bandgap energy of 4.25 eV (Table 2). The obtained value was found to be in agreement with the data reported by Sangappa et al. for regenerated silk fibroin upon γ-irradiation [33]. Incorporation of zinc oxide and copper oxide nanoparticles in the silk fibroin films showed only a decrease of 0.10 eV, in the

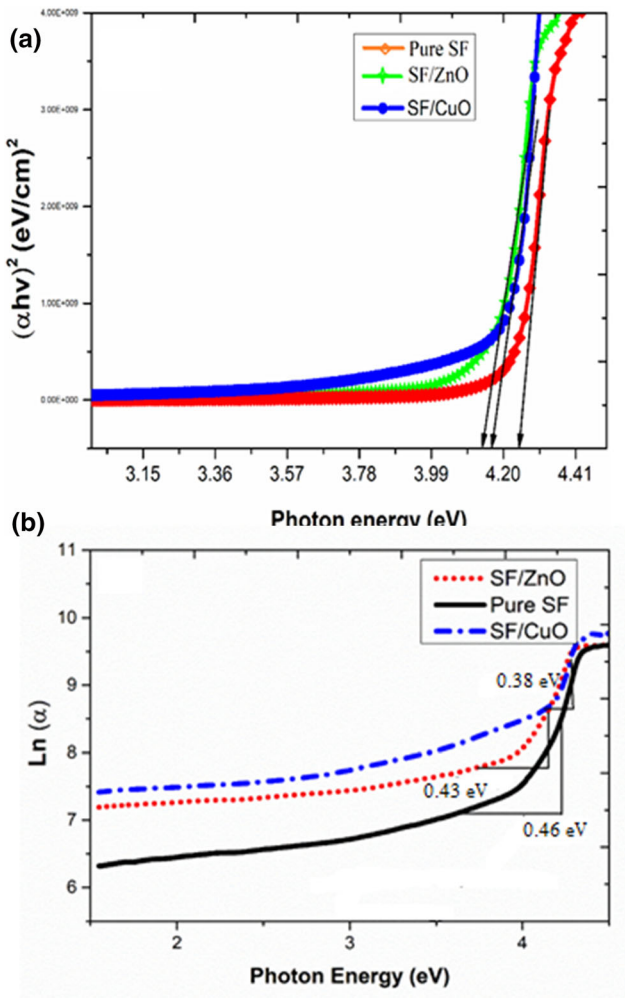


Fig. 4 **a** Optical bandgap energy **b** Urbach energy of pure silk fibroin film, SF/ZnO (0.75) and SF/CuO (0.75) nanocomposite films (thickness 30 μm)

bandgap energy. This observation is suggestive of some possible structural alterations in nanocomposite films.

Urbach energy (E_U) is governed by the structural disorder, an imperfection in stoichiometric and passivation at the surface. It indicates the disorder of photon states of the film in the UV–visible region. Urbach energy has been found to be below the absorption band edge of films and it characterizes the degree of the absorption edge smearing due to the crystalline lattice disordering caused by structural peculiarities [33, 38]. Urbach energy is determined by taking the reciprocal of the slope of the line obtained by plotting the logarithm of absorption coefficient (ln α) against photon energy (hv) using Eq. 6 and presented in Fig. 4b. It was observed that incorporation of zinc oxide nanoparticles does not affect the Urbach energy, however, the addition of copper oxide in the silk fibroin matrix affects the Urbach energy of nanocomposite films (Fig. 4b). Urbach energy of pure SF and SF/ZnO and SF/CuO nanocomposite films are mentioned in Table 2. Change in the band tail feature (Urbach energy) points out the role of metal oxide nanoparticles within the silk fibroin matrix. The decrement in Urbach energy may be attributed to lowering of crystallinity in nanocomposite films [5, 33].

One of the most important parameters for the design of optical devices is the refractive index (n). In this analysis, refractive index was calculated from the absorption coefficient of the films at 550 nm wavelength using Eqs. 7 and 8 ("Electrical and optical properties of silk fibroin/metal oxides nanocomposite films" section) and values are mentioned in Table 2. The refractive index of pure SF film is found 1.61. The refractive index values reported in the literature for regenerated silk films are 1.54 and 1.52 [39, 40]. The addition of metal oxide nanoparticles reduces the refractive index of nanocomposite films. Decrease in the refractive index of nanocomposite films can be related to the β sheet content, crystallinity, and water content of these films [41]. Decrease in refractive index of SF/ZnO and SF/CuO nanocomposite film also reflects the effect of packing density of the metal oxides in silk fibroin. Yusop et al. have shown that as the packing density of ZnO increases in PVC/ZnO nanocomposite films, refractive index increases [42].

Table 2 Optical parameters i.e. bandgap energy, Urbach energy, and refractive index of pure SF and SF/ZnO and SF/CuO nanocomposite films

Sample	Bandgap energy (eV)	Urbach energy (eV)	Refractive index (n) @550 nm
Pure SF	4.25 ± 0.06	0.46 ± 0.07	1.61 ± 0.12
SF/ZnO (0.75)	4.16 ± 0.17	0.43 ± 0.05	1.52 ± 0.08
SF/CuO (0.75)	4.14 ± 0.12	0.38 ± 0.08	1.45 ± 0.14

3.3 Morphology of silk fibroin/metal oxides nanocomposite films

To analyse the homogeneity and morphology of the nanocomposite films, FESEM-EDS studies were carried out. It was observed that pure silk fibroin (pure SF) film had a rough surface because of regeneration and surface etching due to methanol treatment (Fig. 5a). The regenerated silk fibroin (pure SF) films contain α-helix or random coil structure and do not have enough intramolecular bonds to form a stable structure. Such films are water soluble and have low mechanical properties. To enhance the mechanical properties, films are crystallized in organic solvents, to induce a β-sheet structure. In this study, methanol was used as an organic solvent for the coagulation process of the silk fibroin films to induce crystallization. The immersion of films in methanol solution 24 h led to surface etching and rough surface is appeared on the regenerated films [29]. SF/ZnO and SF/CuO nanocomposite films showed uniformly dispersed nanoparticles in the silk fibroin matrix (Fig. 5b, c). However, in the case of SF/CuO film, some agglomeration of CuO nanoparticles was observed and the size of agglomerates was found to be lower than 100 nm. The altered morphology of films indicates the formation of nanocomposites with a homogeneous surface.

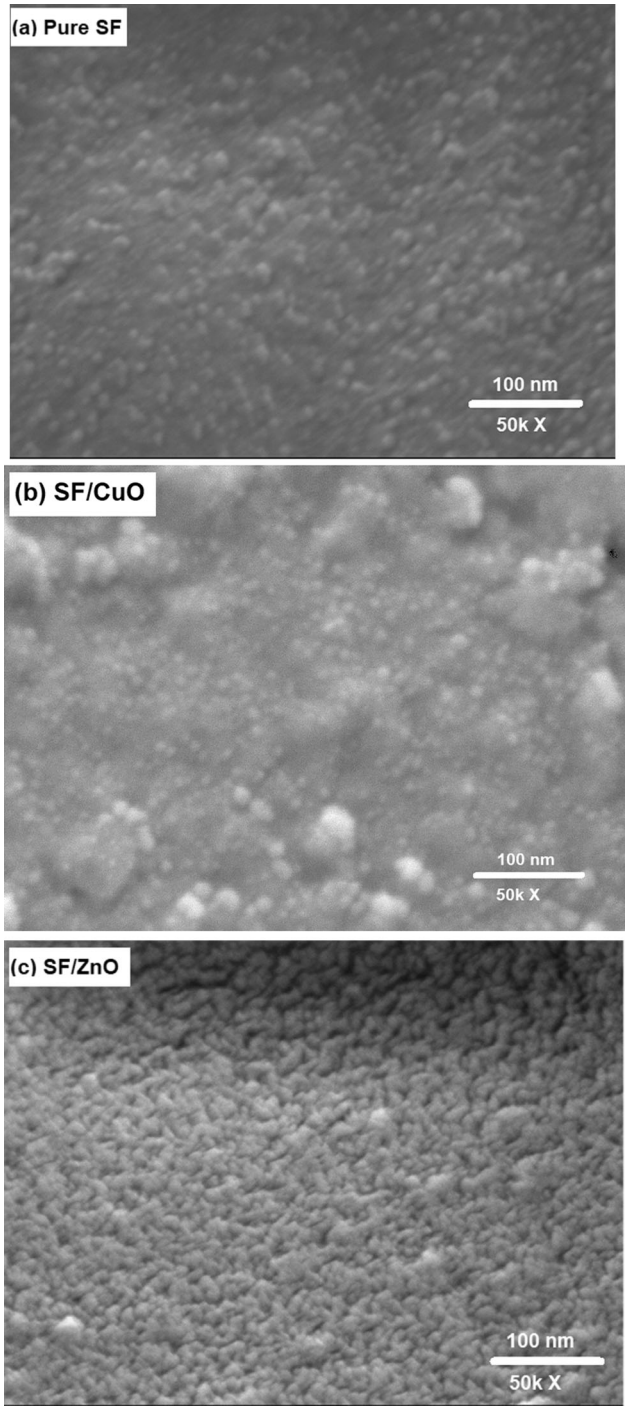


Fig. 5 FESEM micrograph of pure SF and SF/ZnO and SF/CuO nanocomposite films at 50 k X resolution. **a** Pure SF, **b** SF/CuO (0.75), **c** SF/ZnO (0.75)

3.4 Structural properties of silk fibroin/metal oxides nanocomposite films

The FTIR spectroscopy technique is used to determine the secondary structure of the pure SF, SF/ZnO, and SF/CuO nanocomposite films. Silk protein undergoes a conformation transition from random coil, helical conformation, or both to β-sheet when treated with aqueous alcohol or salt solutions [15]. In this study, 80% methanol aqueous solution was used to induce the conformation transition in pure silk fibroin and nanocomposite films. *Bombyx mori* silk protein has three major absorption bands of Amide I (1620–1650 cm^{-1}), Amide II (1520–1540 cm^{-1}), and Amide III (1200–1350 cm^{-1}) region in which amide I and amide III contribute to the β-sheet (silk II) structure. Figure 6 shows the FTIR-ATR spectra of pure SF, SF/ZnO (0.75), and SF/CuO (0.75) nanocomposite films of 30 μm thickness. In the spectra (Fig. 6) it was observed that peak of β-sheet (1620–1630 cm^{-1}) in the Amide I region were overlapped by the peaks of random coil, helical conformation (1655–1660 cm^{-1}), and β turns conformation (1690–1700 cm^{-1}) [34, 43]. So, to compare the secondary structure and β-sheet content, the deconvolution method was applied in the Amide I region,

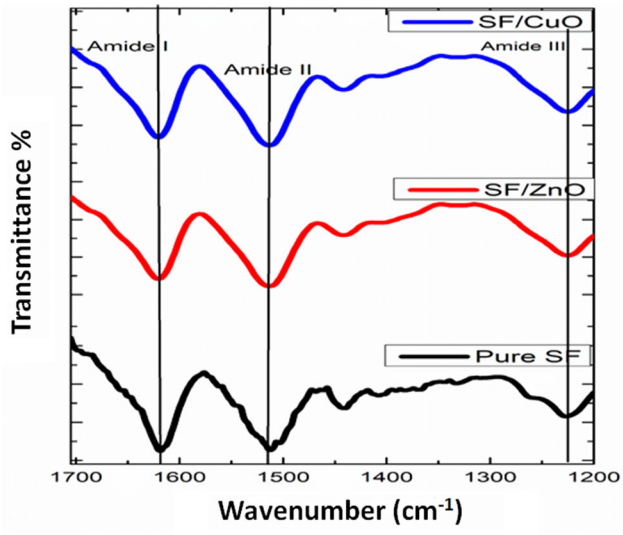


Fig. 6 FTIR-ATR spectra of pure silk fibroin (SF), SF/ZnO (0.75) and SF/CuO (0.75) nanocomposite films with 30 μm thickness

and the graphs are shown in Fig. 7. In the Amide I region of pure SF film (Fig. 7a), the peak at 1616 cm^{-1} is assigned as β-sheet conformation as reported in the literature [6, 43–45]. Similar peaks are observed for SF/ZnO (0.75) and SF/CuO (0.75) nanocomposite films (Fig. 7b, c). The percentage of β-sheet content for pure SF and nanocomposite (SF/ZnO and SF/CuO) films was calculated and values are given in Table 3. The β-sheet content of nanocomposite films is found to be lower than pure SF films which shows that nanocomposite films are more amorphous. These data suggest that metal oxide nanoparticles might behave as a chaotropic agent in the silk fibroin which disturbed its ordered structure. These changes in beta-sheet crystalline structures of silk fibroin due to incorporation of metal oxide nanoparticles explain the observed modification in electrical conductivity, bandgap energy, Urbach energy, and refractive index discussed in the previous sections.

Further, the conformational differences in silk fibroin films were studied by XRD measurement and recorded in the 2-theta range from 10° to 40° and the profiles are shown in Fig. 8. The Diffractogram of pure SF film has the main diffraction peak at $2\theta = 20.6°$ which indicates the presence of β sheet structure in the film [43]. Methanol treatment and annealing process involve the bound water plasticization that promotes β sheet formation via helix–helix interactions [46]. The small peaks at $2\theta = 16.08°$, 19.78°, 20.6°, 23.1°, 24.9°, 25.8°, 27.8°, 32° are observed in pure silk fibroin. These peaks represent the mixture of silk I and silk II structures in the pure SF films. Silk fibroin has two structures known as Silk I ($2\theta = 9°$, 11.9°, 20°, 24.9°, 27.8°, and 32°) and Silk II ($2\theta = 20.6°$ and 25.8°) which represent the amorphous and β turn and β sheet structure respectively [43]. Similarly, SF/CuO nanocomposite films have shown a mixture of silk I and silk II structure. However, the intensity of the signal is lesser than the pure SF film. In SF/ZnO nanocomposite film, no prominent peak was observed indicating the complete amorphous structure. The crystallinity index of the pure SF and SF/ZnO and SF/CuO nanocomposite films was estimated utilizing the area fractions of amorphous and crystalline (i.e. silk I or silk II) peaks in the diffractogram.

To analyse the data, a multi-peak fit program was applied using the Origin pro 9.0 software which separates the overlapped peaks using the standard method (Fig. 9). The crystallinity of the nanocomposite films was estimated by subtracting the area under the amorphous peak from the total integrated area of the XRD pattern, the remaining fraction is assigned to the crystalline phases of silk fibroin using Herman's formula as a similar method is adopted by Naik et al. [43]. Briefly, the degree of crystallinity (X_c) of the polymers or polymer nanocomposites was calculated from the ratio of areas under the crystalline peaks and using the relation [47]:

$$X_c(\%) = \frac{A_c}{(A_c + A_a)} \times 100, \tag{9}$$

where A_c and A_a are the areas of crystalline and amorphous regions in the polymeric samples, respectively. The calculated degree of crystallinity is tabulated in Table 4. It was observed that the degree of crystallinity reduced with the addition of zinc oxide and copper oxide nanoparticles in the silk fibroin films which is consistent with FTIR data. The average crystallite size was calculated by using the Debye–Scherrer equation [8]:

$$P = \frac{k\lambda}{\beta \cos\theta}, \tag{10}$$

where P is the average crystallite size (nm), λ is the wavelength of the X-rays used and its value is 1.5406 Å, k is the Scherrer constant ($k = 0.9$) related to the shape of the crystallite. Here, β is the broadening of the diffraction peak ($2\theta = 20.6°$) measured at full width at half maximum (FWHM) and θ is the diffraction angle. The data of average crystallite size

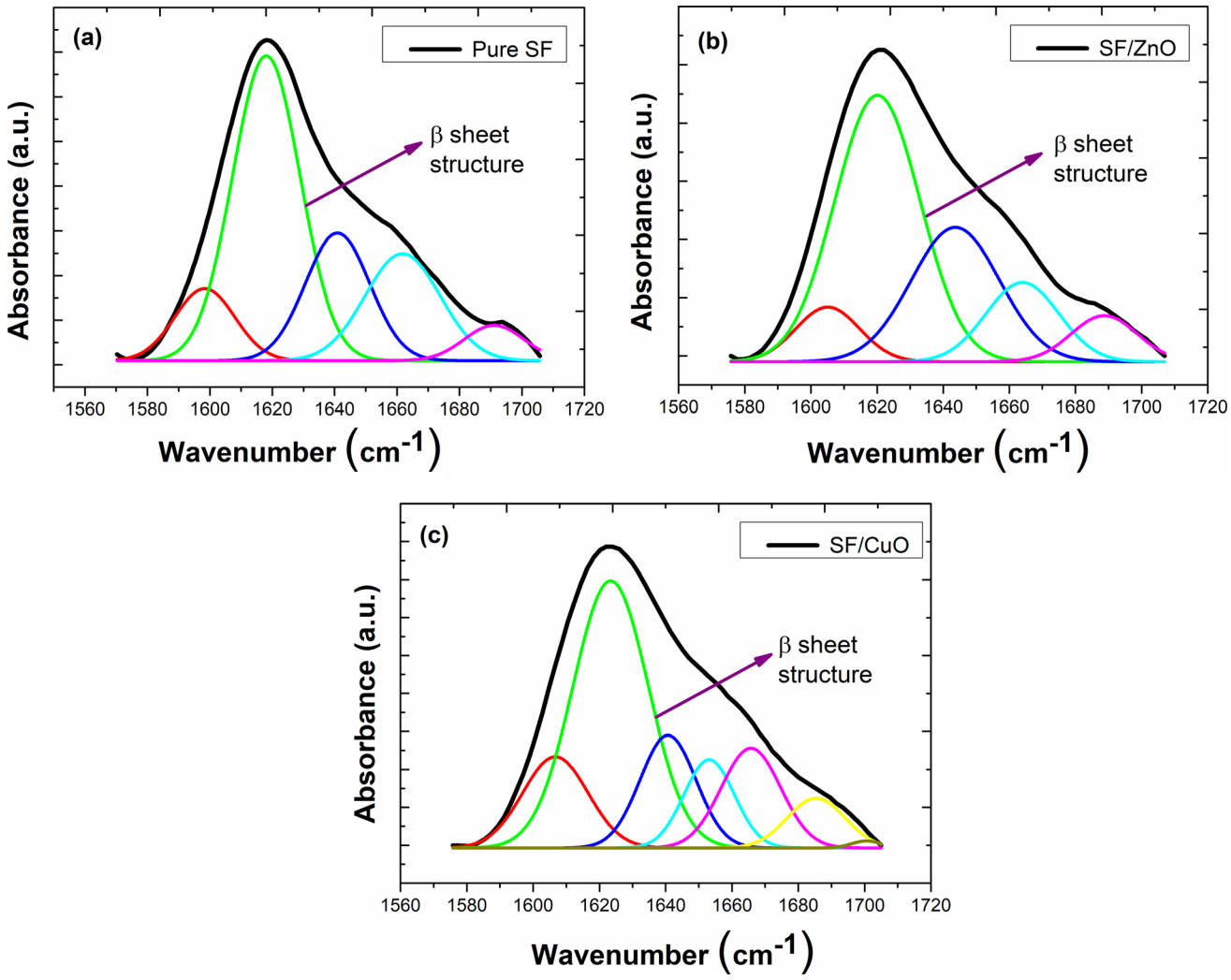


Fig. 7 Deconvolution results of amide I region of silk fibroin films **a** pure SF, **b** SF/ZnO (0.75), **c** SF/CuO (0.75)

Table 3 Comparison of β-sheet content % of pure SF film with the nanocomposite film i.e. SF/ZnO and SF/CuO

Sample	β-sheet content (%)
Pure SF	48.4 ± 0.6
SF/ZnO (0.75)	40.3 ± 0.8
SF/CuO (0.75)	44.5 ± 0.4

are mentioned in Table 4. It can be noticed that the addition of metal oxide nanoparticles in silk fibroin reduced the average crystallite size in the nanocomposite films. The reason behind the lowering of crystallinity and average crystallite size on the incorporation of metal oxide nanoparticles in silk fibroin might be the high surface area of

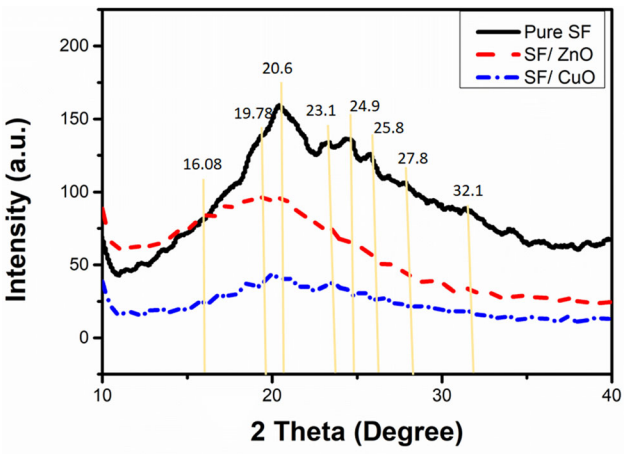


Fig. 8 Diffractogram of pure SF film and SF/ZnO (0.75) and SF/CuO (0.75) nanocomposite films from 2θ 10°–40° measured by XRD

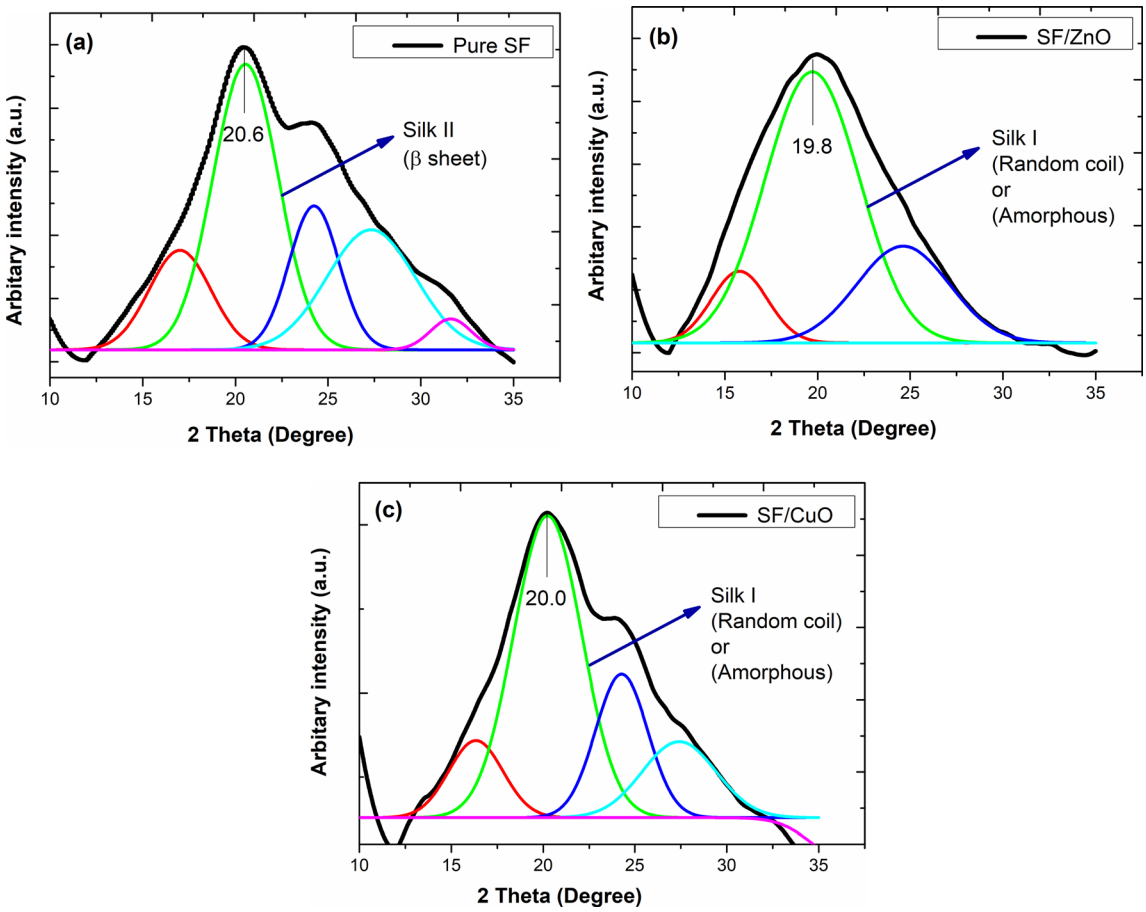


Fig. 9 Multi-peak fit diagram of **a** Pure SF, **b** SF/ZnO (0.75), and **c** SF/CuO (0.75) nanocomposite films from 10 to 35° 2θ angle denoting the silk I and silk II structures peak measured by XRD

Table 4 Structural parameter of pure SF and nanocomposites films i.e. SF/ZnO and SF/CuO

Sample	Crystallinity (%)	Average crystallite size (*P*) (nm)
Pure SF	44.7 ± 1.4	1.5 ± 0.3
SF/ZnO (0.75)	31.1 ± 2.7	0.8 ± 0.5
SF/CuO (0.75)	37.5 ± 1.5	1.2 ± 0.4

nanoparticles affecting the self-assembly of silk fibroin molecules by obstructing the motion of protein chains which results in the nanoconfined crystallization of silk fibroin and hence small crystallite size. The reduced crystallinity and average crystallite size in silk fibroin on incorporation of nanoparticles found in accordance with Zang et al. for SF/GO and SF/TiO_2 nanocomposites [48, 49].

4 Conclusions

The influence of metal oxide nanoparticles on the silk fibroin matrix in terms of conductivity was evaluated. It showed a linear relationship between the concentration of metal oxides and conductivity. Addition of both, ZnO and CuO nanoparticles enhanced the conductivity of silk fibroin nanocomposite films. The SF/ZnO film showed a 1000-fold increment compared to pure silk fibroin film. The conductivity of the films reduced with increase in thickness. SF/CuO film showed a fourfold increment in conductivity while SF/ZnO film showed a tenfold

increment with decrease in thickness. Incorporation of metal oxide nanoparticles in silk fibroin increased the carrier (electrons or holes) concentration examined by Hall Effect. In SF/CuO film, carrier concentration was twice as in pure SF while SF/ZnO film had a 100-fold increment than pure SF film. Positive hall voltage in pure SF and SF/CuO film suggested the p-type semiconducting behaviour while negative hall voltage in SF/ZnO film was concluded as n-type semiconducting behaviour. Optical transmittance of silk fibroin/metal oxides nanocomposites was found to be higher than pure Silk fibroin film. Addition of ZnO and CuO nanoparticles in the silk fibroin matrix reduced the Urbach energy and refractive index of nanocomposite films. ZnO and CuO nanoparticles were uniformly dispersed in the silk fibroin matrix. Structural properties of silk fibroin/metal oxides nanocomposite films revealed that addition of metal oxide nanoparticles reduced the β sheet content, crystal size, and crystallinity. The expansion of the amorphous phase in silk fibroin/metal oxides nanocomposite films facilitates the charge mobility and optical transparency. These transparent, metal oxides-incorporated nanocomposite silk fibroin films with enhanced conductivity hold the potential for utilization in numerous optoelectronic devices.

Acknowledgements

The authors thank the financial support by Science & Engineering Research Board (SERB), Department of Science and Technology (DST), Govt. of India (EMR/2017/002833).

References

1. Q. Sun, B. Qian, K. Uto, J. Chen, X. Liu, T. Minari, Biosens. Bioelectron. **119**, 237–251 (2018)
2. L.D. Koh, J. Yeo, Y.Y. Lee, Q. Ong, M. Han, B.C.K. Tee, Mater. Sci. Eng. C **86**, 151–172 (2018)
3. C. Wang, K. Xia, Y. Zhang, D.L. Kaplan, Acc. Chem. Res. **52**, 2916–2927 (2019)
4. S.N. Fan, Y. Zhang, X.Y. Huang, L.H. Geng, H.L. Shao, X.C. Hu, Y.P. Zhang, Sci. China Technol. Sci. **62**, 903–918 (2019)
5. R. Ranjana, N. Parushuram, K.S. Harisha, S. Asha, B. Narayana, M. Mahendra, Y. Sangappa, J. Mater. Sci. Mater. Electron. **31**, 249–264 (2020)
6. M.H. Park, J. Kim, S.C. Lee, S.Y. Cho, N.R. Kim, B. Kang, E. Song, K. Cho, H.J. Jin, W.H. Lee, RSC Adv. **6**, 5904–5914 (2016)
7. R. Jung, H.S. Kim, Y. Kim, S.M. Kwon, H.S. Lee, H.J. Jin, J. Polym. Sci. B **46**, 1235–1242 (2008)
8. C.S. Shivananda, B. Lakshmeesha Rao, Y. Sangappa, J. Mater. Sci. Mater. Electron. **31**, 41–51 (2020)
9. T. Posati, V. Benfenati, A. Sagnella, A. Pistone, M. Nocchetti, A. Donnadio, G. Ruani, R. Zamboni, M. Muccini, Biomacromol **15**, 158–168 (2014)
10. H. Zhang, J. Zhao, T. Xing, S. Lu, G. Chen, Polymers **11**, 1774 (2019). https://doi.org/10.3390/polym11111774
11. S.H. Tsao, D. Wan, Y.S. Lai, H.M. Chang, C.C. Yu, K. Te Lin, H.L. Chen, ACS Nano **9**, 12045–12059 (2015)
12. M. Kang, H.J. Jin, Colloid Polym. Sci. **285**, 1163–1167 (2007)
13. D. Naskar, P. Bhattacharjee, A.K. Ghosh, M. Mandal, S.C. Kundu, MACS Appl. Mater. Interfaces **9**, 19356–19370 (2017)
14. J. Ayutsede, M. Gandhi, S. Sukiraga, H. Ye, C.M. Hsu, Y. Gogotsi, F. Ko, Biomacromol **7**, 208–214 (2006)
15. Y. Wang, R. Ma, K. Hu, S. Kim, G. Fang, Z. Shao, V.V. Tsukruk, A.C.S. Appl, Mater. Interfaces **8**, 24962–24973 (2016)
16. T. Cohen-Karni, K.J. Jeong, J.H. Tsui, G. Reznor, M. Mustata, M. Wanunu, A. Graham, C. Marks, D.C. Bell, R. Langer, D.S. Kohane, Nano Lett. **12**, 5403–5406 (2012)
17. M. Jia, N. Qi, Q. Dong, H. Wang, K.-Q. Zhang, M.L. Qi, N. Dong, Q. Wang, H. Zhang, Int. J. Nanotechnol. **11**, 298–310 (2014)
18. K. Min, M. Umar, H. Seo, J.H. Yim, D.G. Kam, H. Jeon, S. Lee, S. Kim, RSC Adv. **7**, 574–580 (2017)
19. S.R.V. Siva Prasanna, K. Balaji, S. Pandey, S. Rana, *Nanomaterials and polymer nanocomposites—raw materials to applications* (Elsevier, London, 2018), pp. 123–144
20. S. Vallabhapurapu, L.D. Varma Sangani, M. Ghanashyam Krishna, J. Das, C. Tu, S. Du, A. Srinivasan, Acta Phys. Pol. A **66**, 68–70 (2018)
21. D.M. Jundale, S.T. Navale, G.D. Khuspe, D.S. Dalavi, P.S. Patil, V.B. Patil, J. Mater. Sci. Mater. Electron. **24**, 3526–3535 (2013)
22. K. Zilberberg, J. Meyer, T. Riedl, J. Mater. Chem. C **1**, 4796–4815 (2013)
23. K. Vidhya, M. Saravanan, G. Bhoopathi, V.P. Devarajan, S. Subanya, Appl. Nanosci. **5**, 235–243 (2015)
24. Y. Cheng, G. Yu, Z. Duan, J. Nanomater. **2019**, 9481415 (2019). https://doi.org/10.1155/2019/9481415
25. M. Dolores Bermúdez, W. Brostow, F.J. Carrión-Vilches, J. Sanes, J. Nanosci. Nanotechnol. **10**, 6683–6689 (2010)

26. D.H. Xie, F.F. Wang, H. Lü, M.Y. Du, W.J. Xu, Chin. Phys. B **22**, 58103 (2013). https://doi.org/10.1088/1674-1056/22/5/058103
27. A. Manjunath, M. Irfan, K.P. Anushree, K.M. Vinutha, N. Yamunarani, Adv. Mater. Phys. Chem. **06**, 263–273 (2016)
28. A. Arena, G. Scandurra, C. Ciofi, Sensors **17**, 2198 (2017). https://doi.org/10.3390/s17102198
29. C.M. Srivastava, R. Purwar, R. Kannaujia, D. Sharma, Fibers Polym. **16**, 1020–1030 (2015)
30. S. Lu, X. Wang, Q. Lu, X. Zhang, J.A. Kluge, N. Uppal, F. Omenetto, D.L. Kaplan, Biomacromol **11**, 143–150 (2010)
31. H. Chandra, S.W. Allen, S.W. Oberloier, N. Bihari, J. Gwamuri, J.M. Pearce, Materials **10**, 110 (2017). https://doi.org/10.3390/ma10020110
32. O. Bierwagen, T. Ive, C.G. Van De Walle, J.S. Speck, Appl. Phys. Lett. **93**, 242108 (2008). https://doi.org/10.1063/1.3052930
33. R. Madhukumar, S. Asha, B. Lakshmeesha-Rao, B.K. Sarojini, K. Byrappa, Y. Wang, Y. Sangappa, Radiat. Eff. Defects Solids **170**, 906–915 (2015)
34. S. Ling, Z. Qi, D.P. Knight, Z. Shao, X. Chen, Biomacromol **12**, 3344–3349 (2011)
35. T. Liang, Y. Makita, S. Kimura, Polymer **42**, 4867–4872 (2001)
36. R.K. Sonker, S.R. Sabhajeet, B.C. Yadav, J. Mater. Sci. Mater. Electron. **27**, 11726–11732 (2016)
37. A. Bucciarelli, V. Mulloni, D. Maniglio, R.K. Pal, V.K. Yadavalli, A. Motta, A. Quaranta, Opt. Mater. **78**, 407–414 (2018)
38. E.R. Shaaban, N. Afify, A. El-Taher, J. Alloys Compd. **482**, 400–404 (2009)
39. M.B. Applegate, G. Perotto, D.L. Kaplan, F.G. Omenetto, Biomed. Opt. Express **6**, 4221 (2015). https://doi.org/10.1364/boe.6.004221
40. S.T. Parker, P. Domachuk, J. Amsden, J. Bressner, J.A. Lewis, D.L. Kaplan, F.C. Omenetto, Adv. Mater. **21**, 2411–2415 (2009)
41. G. Perotto, Y. Zhang, D. Naskar, N. Patel, D.L. Kaplan, S.C. Kundu, F.G. Omenetto, Appl. Phys. Lett. **111**, 103702 (2017). https://doi.org/10.1063/1.4998950
42. W. Al-Taa'Y, M. Abdul Nabi, R.M. Yusop, E. Yousif, B.M. Abdullah, J. Salimon, N. Salih, S.I. Zubairi, Int. J. Polym. Sci. **2014**, 697809 (2014). https://doi.org/10.1155/2014/697809
43. M.B. Dickerson, S.P. Fillery, H. Koerner, K.M. Singh, K. Martinick, L.F. Drummy, M.F. Durstock, R.A. Vaia, F.G. Omenetto, D.L. Kaplan, R.R. Naik, Biomacromol **14**, 3509–3514 (2013)
44. X. Hu, K. Shmelev, L. Sun, E.S. Gil, S.H. Park, P. Cebe, D.L. Kaplan, Biomacromol **12**, 1686–1696 (2011)
45. I.C. Um, H.Y. Kweon, K.G. Lee, Y.H. Park, Int. J. Biol. Macromol. **33**, 203–213 (2003)
46. K. Yazawa, K. Ishida, H. Masunaga, T. Hikima, K. Numata, Biomacromol **17**, 1057–1066 (2016)
47. P.H. Hermans, A. Weidinger, J. Appl. Phys. **19**, 491–506 (1948)
48. C. Zhang, Y. Zhang, H. Shao, X. Hu, ACS Appl. Mater. Interfaces **8**, 3349–3358 (2016)
49. H. Pan, Y. Zhang, H. Shao, X. Hu, X. Li, F. Tian, J. Wang, J. Mater. Chem. B **2**, 1408–1414 (2014)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

2020 IEEE VLSI Device, Circuit and System Conference (VLSI-DCS), 18-19 July, 2020, ED MSIT SBC, Kolkata, India

TCAD Analysis and Modelling of Gate-Stack Gate All Around Junctionless Silicon NWFET Based Bio-Sensor for Biomedical Application

Mekonnen Getnet Yirak
Department of Applied Physics
Delhi Technological University
Delhi-110042, INDIA
mekonnengetnet01@gmail.com

Rishu Chaujar
Department of Applied Physics
Delhi Technological University
Delhi-110042, INDIA
chaujar.rishu@dtu.ac.in

Abstract--**In the present day, metallic oxide semiconductor field-effect transistor-based bio-sensors have been frequently used for various purposes due to their low cost and other properties. In this work, high-k Gate-Stack gate-all-around junctionless Silicon Nanowire FET (SiNWFET) is proposed for neutral biomolecule species detection and enhanced the device performance by introduced gate stack and high metal gate work-function. In particular, neutral biomolecule species like Streptavidin, Uricase, APTES, Protein and ChO_x are considered in our study. Subthreshold slope, drain induced barrier lowering (DIBL), leakage current, transconductance, and shifting threshold voltage were considered for study the bio-sensor response. Effect of cavity thickness, cavity length, High-k dielectric thickness, and its length on the detection of the device has also become examined. The results in gate stack junctionless gate all around SiNWFET shows better performance in terms of DIBL, transconductance, leakage current, I_{ON}/I_{OFF} ratio and subthreshold slope. The high-k dielectric oxide (HfO_2) has been identified for chemical compatibility and thermal stability properties on metal oxide semiconductor transistor as a gate oxide to mitigate the gate tunneling current and short channel effects.**

Keywords: - ATLAS-3D, Bio-sensor, Gate-Stack Gate-all- around (GS-GAA), Junctionless, SiNWFET, Neutral-biomolecule

I. INTRODUCTION

The integrated circuit industry has shown a lot of technological progress since the past years. Metal Oxide Semiconductor Field Effect Transistor (MOSFET) is a fundamental component of the Integrated Circuit (IC)[1]. The integrated circuit had advanced with increasing the number of on-chip transistors every generation and reducing the physical device size[2]. It has attracted many possibilities of novel structures in MOSFET devices.

With the development of Moore's law, the size of the transistor is decreasing so fast, led to various problems like mechanical tunneling, short channel effect (SCEs) and drain induced barrier lowering effect will be prominently related to a reduction of device size as channel length comparable to bandwidth [3][4]. It causes body-bias and electron tunneling into the gate or body/substrate and generally degrades device performance [5].

To overcome doping profile complexity, Junctionless MOSFET (JLMOSFET) without any pn-junction in the source-drain channel path is used for tiny channel length and thickness transistors [6]. It has a homogeneous and uniform doping profile throughout the source- channel-drain, unlike a junction-based (JB) inversion mode (IM) transistor[6][7].

In addition to Junctionless MOSFET, several devices related to different geometry have been proposed and validated experimentally [6][8]. Nowadays, junctionless multi-gate device structures are used due to their scaling capability, enhancing gate controllability, low cost, reducing short channel effects, and more controlling power over the channel region[10].

From many multi-gate MOSFET devices structures, Cylindrical Gate All Around Transistor device has preferred; due to its superior electrostatic control ability over the channel and transport property. In this device, the silicon channel is wholly surrounded by the metal gate to suppress short channel effects[11]. It is considered as one of the best multi-gate structures as it shows better gate controllability and excellent CMOS compatibility[12] [4].

Apart from device miniaturization, junctionless gate all around transistor has an increasing demand of establishing conventional clinical laboratories in rural, remote areas and military fields for the detection of biomolecular species [13] [14] and also overcome problems like short channel effects, hot-electron, and threshold voltage roll-off[15][16].

Reducing gate dielectric oxide thickness (< 2nm) increases mechanical tunneling and short channel effects, which can degrade the circuit operation and device performance [17][18]. Also, continually decrease gate oxide thickness led to direct tunneling[19] and excessive leakage currents in Si-based MOS Devices. The material hafnium Oxide (HfO_2) of the high-k dielectric family has better thermal stability, and large energy bandgap (5.8ev) as compared to silicon (1.1ev) can reduce problems related to short channel effect[20] and HfO_2 is also known as a better electrical insulator, and it is a stable, inorganic colorless solid compound[19][20]. Using HfO_2 in MOS device Modeling as gate Oxide, results in a better gate control over SiO_2[20].

In this paper, we have proposed, n-type junctionless gate stack gate all around silicon nanowire-based biosensor transistor using the method of gate material engineering (GME) at reduced gate oxide thickness (<2.0) using high-k material (HfO_2) as gate oxide above silicon dioxide layer with the metal gate having high work-functions to realize the impact of neutral biomolecule dielectric constant variation on transfer characteristics of the proposed device, such as I_{ON}/I_{OFF} ratio,

978-1-7281-1933-5/20/$31.00 ©2020 IEEE
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:35:54 UTC from IEEE Xplore. Restrictions apply.

shifting threshold voltage, and drain-current sensitivity of GS-GAAJLT for the first time.

II. DEVICE STRUCTURE AND SIMULATIONS

Table 1. Device parameter and structure

Parameters	GAAJLT	GS-GAAJLT
Channel length (nm)	20	20
Work function (eV)	4.8	4.8
Gate Oxide thickness (nm)	SiO_2=1.8	SiO_2=0.3 & HfO_2=1.5
Diameter of silicon film (nm)	10	10
Source/Drain & Channel Doping (N_D++)	$10^{19}cm^{-3}$	$10^{19}cm^{-3}$
Oxide Dielectric constant	SiO_2=3.9	HfO_2=20.0

Based on the above device specification, numerous simulations were carried out using atlas-3D device simulator tool for both GAA-JL-SiNWFET and Gate Stack GAA JL-SiNWFET structure. "Fig.1", shows (a) 3D-schematic and (a) 2D cross-sectional view of gate stack junctionless gate all around transistor with silicon nanowire channel thickness (t_{si}=10nm) and channel length (L= 20 nm). Source and drain lengths are considered to be 10 nm long. Both oxide thicknesses are 1.8nm (t_{ox} = 0.3nm + thigh-k = 1.5nm). Length of a metal Gate denoted by l_1 and L_2 represents SiO_2 and cavity lengths near to the metal gate, respectively, their sum equal to the channel length. In this study, neutral biomolecules are simulated by introducing their dielectric constant in the cavity region; such as choli oxidase (ChO_X =3.3), protein=2.5, uricase=1.54, streptavidin=2.1, and APTES=3.57. also, the air cavity is considered by introducing it's dielectric constant, ϵ =1.00 [12][21].

Fermi-Dirac Statics also used for simulation. Doping versus mobility can be demonstrated by concentration-dependent mobility (CONMOB)[22]; this model used to simulate band gap mobility in the high channel doping along the bandgap narrowing (BGN) model. Shockley–Read–Hall (SRH) model along the Boltz-man transport model has been used to account recombination of minority carriers within the drift or diffusion approximation in the semiconductor without impact ionization model. Carrier-carrier scattering mobility model (CCSMOB) has been used at higher carrier concentration. For perpendicular and parallel field-dependent mobility, the CVT mobility model has been used[22].

III. GATE OXIDE THICKNESS FOR GS-GAAJLT

Gate insulator equivalent oxide thickness (EOT) in this study obtained use "(1)", [23].

$$T_{SiO_2}=T_{high\text{-}k}\times\frac{\epsilon_{SiO_2}}{\epsilon_{High\text{-}k}} \qquad (1)$$

So that, 20nm channel length and 10nm diameter silicon film nanowire was warped by 0.3nm and 1.5nm thick Gate oxides, SiO_2 (ϵ=3.9) and HfO_2 (ϵ=20), respectively. Source /drain and silicon channel all are uniformly doped ($10^{19}cm^{-3}$) with n-type material/impurity.

A high work-function metal gate electrode (4.8eV) was used here to obtain a necessary threshold voltage, and it suppresses DIBLeffects and parasitic gate resistance by enhancing the drive current and transconductance. Gold (Au) electrode has a low electrical resistivity at 20 °C(2.44 μΩ cm) [24]. Because of these properties, it is often used in MOSFET device modeling as an electrode. So, we have used it as a gate electrode in our device structure.

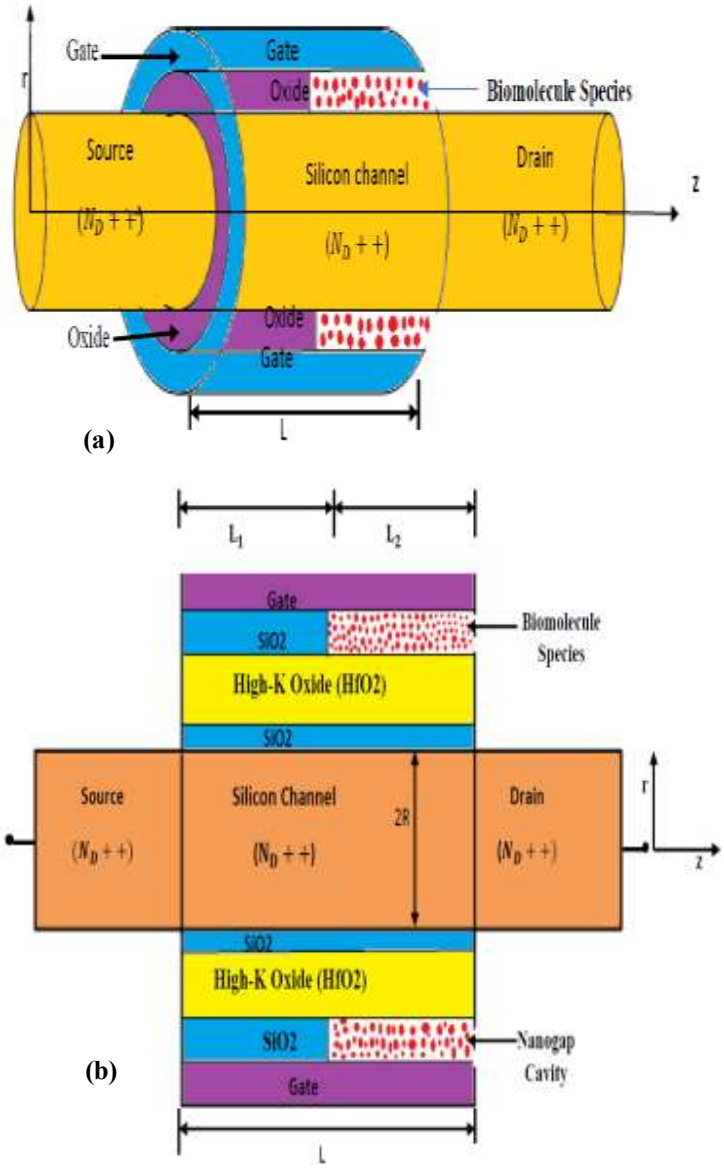


Fig.1 (a) 3D schematic Structure and (b) 2D cross-sectional view with cavity region of gate stack junctionless GAA-SiNWFET.

IV. DEVICE CALIBRATION

To validate our work, the models applied to the device simulation for GAAJLT[11] were applied to GS-GAAJLT. The calibration methods used in this simulation are

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:35:54 UTC from IEEE Xplore. Restrictions apply.

demonstrated in "Fig.2", by changing the dielectric constant in the cavity region, neutral biomolecules can be studied in both GS-GAAJLT and GAAJLT device structure.

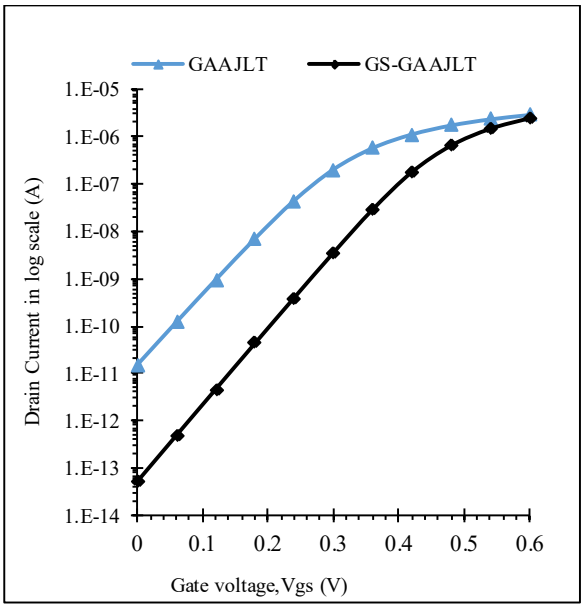


Fig.2. Calibration simulation results of the two devices without a cavity region at V_{DS}= 0.1V.

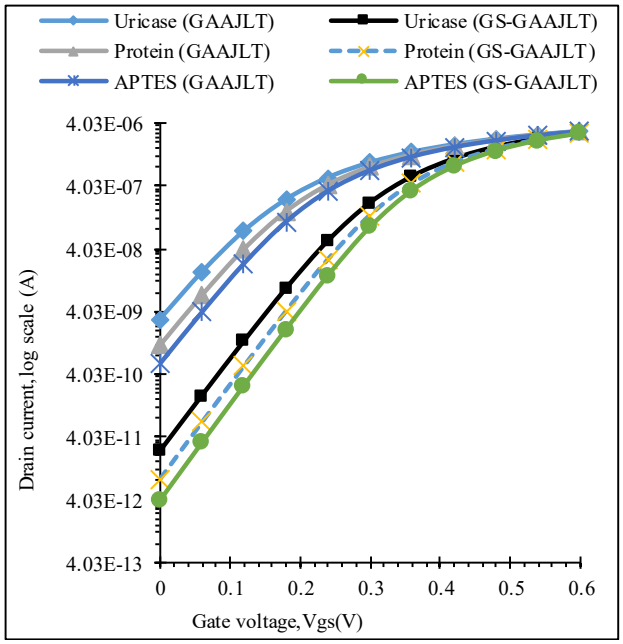


Fig.3. Drain current transfer characteristics of neutral biomolecule species in both devices at V_{DS}= 0.1V.

As we have seen from figure 2 and 3, in both cases I_{OFF} is lower under GS-GAAJLT than GAAJLT. The reason is that hafnium oxide (HfO_2) produces high gate capacitance with a neutral biomolecule, which leads to more drive current, I_D thereby reducing leakage current, DIBL, SCE, and SS in GS-GAA-SiNWFET[25]. Gate dielectrics of SiO_2 is used to make compatible HfO_2 dielectric in the GS-GAAJLT. Since, HfO_2 is a useful tool to improve electron transport efficiency and uniformity of electric field distribution in the channel region[19][26]. Also, HfO_2 suppress the fringing electric field, resulting in an increase in on-state current in GS-GAAJLT[27].

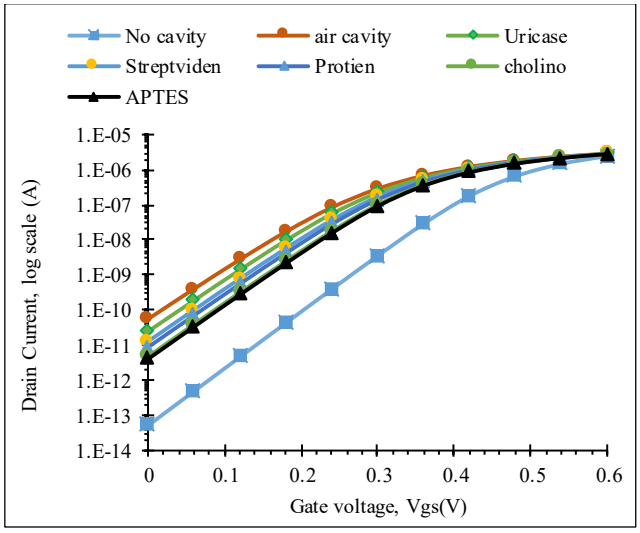


Fig. 4. Drain current transfer characteristics of junctionless gate stack gate all around bio-sensor SiNWFET at V_{DS}= 0.1V.

When a cavity region is filled by target biomolecule dielectric constant, drain ON and OFF current changes are observed. The result shows that, as neutral biomolecule dielectric constant increases, both I_{ON} and I_{OFF} -currents vary. But the shift in OFF-current is more significant than the change in ON-current due to the presence of neutral biomolecules in the nanogap cavity. In this case, OFF-current shows higher sensitivity than ON-current because the shift/change in the dielectric is more effective below the threshold voltage so that the new figure of merit $S_{I_{OFF}}$ use "(3)", has been introduced to study device sensitivity.

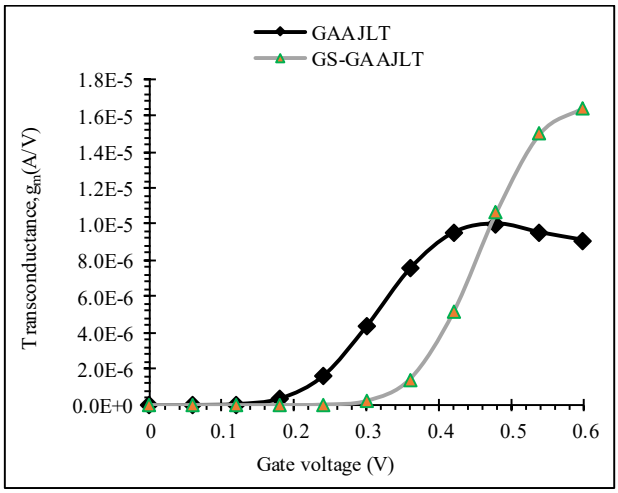


Fig.5. Transconductance variation(g_m) with V_{GS} for both device at V_{DS}=0.1V.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:35:54 UTC from IEEE Xplore. Restrictions apply.

"Fig.5", Demonstrates the comparison of transconductance characteristics in both devices, GS-GAAJLT shows an increase in g_m by 79.63% at V_{DS}=0.1V and V_{GS}=0.6V without cavity region in both devices. The charge density created by HfO_2 in the GS-GAAJLFET shows improved drain current and transconductance[20], thereby reducing hot carrier's effects.

V. Result and Discussion

A) Drain induced barrier lowering

Due to different gate oxide dielectrics and layer structure, the extracted results are different in both devices. High-k gate oxide (HfO_2) in GS-GAAJLT reduces hot-carrier effects, SCEs, and velocity saturation[27]. And also reduces DIBL; it can be obtained use"(2)",[23].

$$DIBL=\frac{\Delta V_{th}}{\Delta V_{DS}} \qquad (2)$$

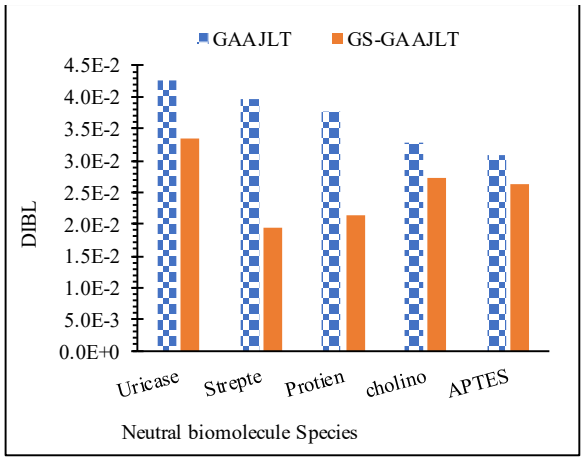


Fig.6. Comparison of DIBL in GS-GAAJLT and GAAJLT at V_{D1} and V_{D2}=0.1V and 3V, respectively at (V_{GS}=0.6V& V_{DS}=0.1V).

B) Subthreshold slope (SS)

The subthreshold slope can be defined as the reciprocal of the slope of the log (I_{DS}) vs. V_{GS}, and it should be small as possible since it determines the amount of voltage swing necessary to switch a device from its "OFF" state to its "ON" state[28]. Fig.7.Illustrates the variation of subthreshold slope for different biomolecule species in both devices. For instance, at APTES, neutral biomolecule SS in GS-GAAJLT is 63mV/decade while 73mV/decade is in GAAJLT. The reason is that mobility (electron and holes) transport efficiency of the GS-GAAJLT device improved by HfO_2[20][21] and reduction of bandwidth and SCEs in the channel.

C) Leakage current

"Fig.7", illustrates variation of leakage current and subthreshold slope for different biomolecule species in both devices. For instance, (without cavity region), the OFF-current is decreased by $3.42x10^{-3}$ times in GS-GAAJLT. That means, extracted values OFF-current in both GAAJLT and GS-GAAGJLT are $1.5x10^{-11}$ and $5.13x10^{-14}$ Ampere respectively. In general, when the neutral biomolecule is present in the cavity region near to HfO_2 gate oxide, gate control is improved incredibly leads to an increase of gate capacitance, reduces threshold voltage roll-off[14], and reduces short channel effects in GS-GAAJLT.

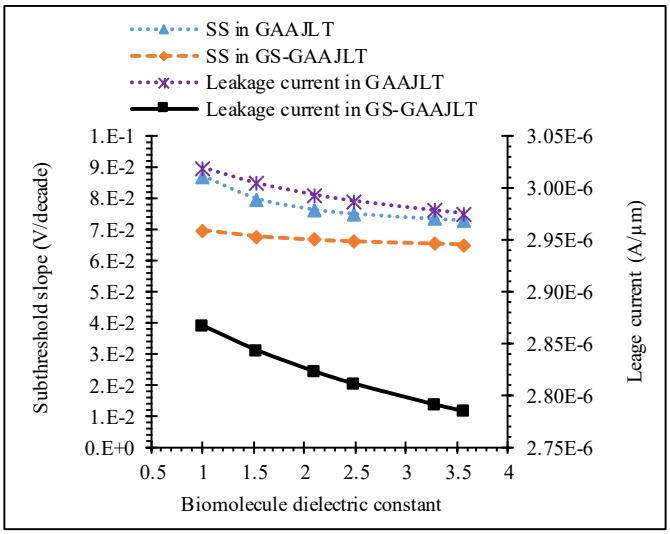


Fig.7. Effect of biomolecule species on subthreshold slope and leakage current for both devices at (V_{GS}=0.6V& V_{DS}=0.1V).

D) Switching (I_{ON}/I_{OFF}) ratio

"Fig.8", demonstrates higher switching (I_{ON}/I_{OFF}) ratio in gate stack gate all around junctionless silicon nanowire transistor. This proves that High-k gate oxide controls short channel effects in the gate stack junctionless gate all around transistor compared to GAAJLT. We examine that, OFF-current decreases in Hafnium oxide gate device than SiO_2 device due to high mobility carries and low bandwidth in HfO_2 gate oxide [19][23] transistor.

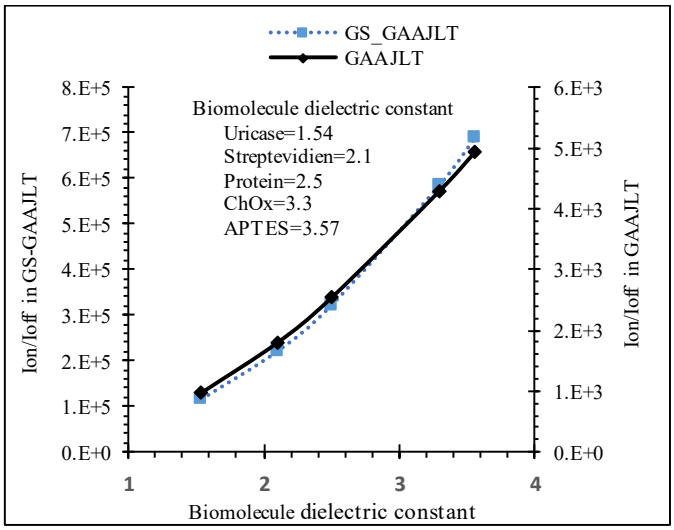


Fig.8. Effect of biomolecule species on switching (I_{ON}/I_{OFF}) ratio on both devices at (V_{GS}=0.6V& V_{DS}=0.1V).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:35:54 UTC from IEEE Xplore. Restrictions apply.

VI. The Sensitivity of Neutral Biomolecule Species

A) Variation of Sensitivity with I_{OFF} ratio ($S_{I_{OFF}}$)

Using different biomolecule dielectric constants in the cavity region, and changing the cavity thickness and length, the sensitivity of the devices was studied. The result shows that $S_{I_{OFF}}$ is increased by 96% in GS-GAAJLT than GAAJLT using APTES neutral biomolecule species in the same condition, this is the effect of hafnium oxide (HfO_2) gate material, and it is more demonstrated by "Fig. 9,10 and 11". Sensitivity ($S_{I_{OFF}}$) of the device is obtained use "(3)" [11],

$$S_{I_{OFF}} = \frac{I_{off}(\text{With Biomolecule Species})}{I_{off}(\text{Without Biomolecule Species})} \Big| \text{at } V_{gs}=0 \qquad (3)$$

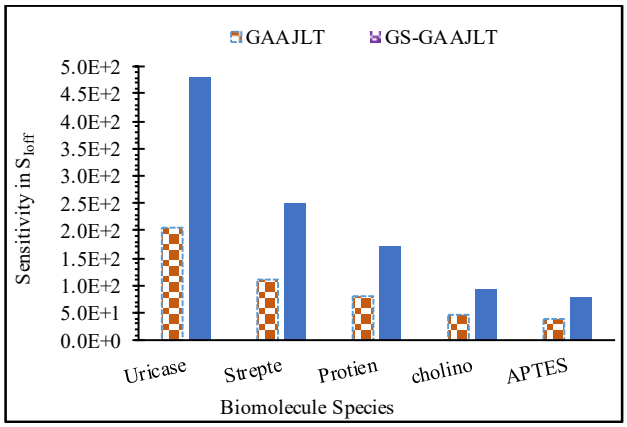


Fig.9. Effect of biomolecule Species on drain-off current sensitivity parameter in both devices at (V_{GS}=0.6V& V_{DS}=0.1V).

Variation of sensitivity is demonstrated by "Fig.12", when the nanogap cavity increases from 0.5nm to 1.5nm; due to large number of neutral biomolecules present in the nanogap cavity region. Increasing the nanogap cavity thickness leads to a higher number of neutral biomolecules to be settled in the nanogap cavity. So, the effect of biomolecules on the device performance is enhanced, and its sensitivity increases as a result of a high-k gate oxide material.

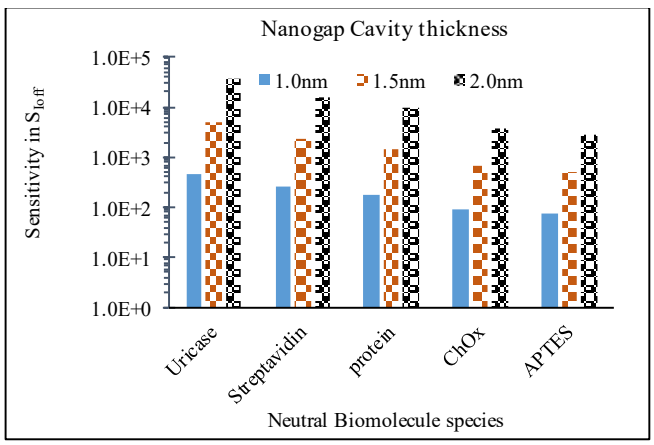


Fig.10. Effect of biomolecule species on drain OFF-current sensitivity parameter in GS-GAA-JLT at (V_{GS}=0.6V& V_{DS}=0.1V).

B) Variation of Sensitivity with shifting threshold Voltage

The ΔV_{th} increases with the biomolecule dielectric constant because the gate voltage exerts higher potential on the channel barrier through the capacitive coupling effect of the neutral biomolecule species and reduces threshold voltage roll-off and mechanical tunneling[29].

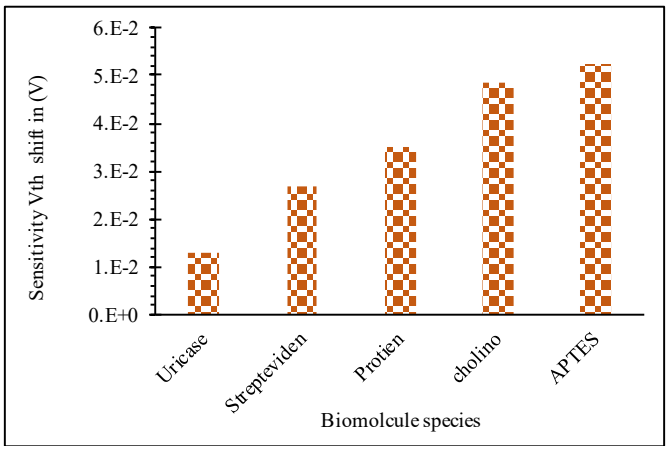


Fig.11. Threshold voltage shifting with neutral biomolecules in the GS-GAA-JLT device at (V_{GS}=0.6V& V_{DS}=0.1V).

"Fig. 11 and 12", demonstrates, shifting threshold voltage depends on high-k dielectric constant, neutral biomolecule dielectric constant, and nanogap cavity thickness. Both threshold voltage shifting and I_{OFF} ratio are used to detect neutral biomolecule species[11][29]. Sensitivity of the device in terms of shifting threshold voltage (ΔV_{th}) can be obtained use "(3)",[29],

$$\Delta V_{th} = |V_{th}(\epsilon_{bio}) - V_{th}(\epsilon_{air})| \qquad (3)$$

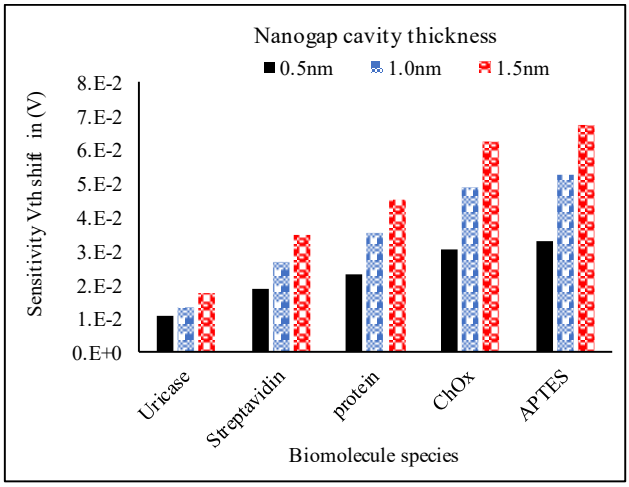


Fig.12.Threshold voltage shifting with biomolecules along with cavity thickness in GS-GAAJLT at (V_{GS}=0.6V& V_{DS}=0.1V).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:35:54 UTC from IEEE Xplore. Restrictions apply.

VII. CONCLUSION

In this paper, the sensitivity of junctionless gate stack gate all around silicon nanowire field-effect transistor for detection of neutral bio-molecules species has been studied. Transfer characteristics such as DIBL, drain current, transconductance, subthreshold slope also examine in the study. The effect of cavity thickness and length of the device also observed. The result shows that electrical characteristics, i.e., subthreshold slope, leakage current, and DIBL, are lower in GS-GAAJLT. Transconductance and switching (I_{ON}/I_{OFF}) ratio shows better results in gate stack gate all around junctionless transistor. So that I_{OFF} ratio has been taken as a sensing metric for the detection of neutral biomolecule species. Sensitivity in terms of I_{OFF} ratio in GS-GAAJLT using APTES neutral biomolecule in both devices in the same condition is improved by 96% than GAAJLT in our study. We can conclude that GS-GAA-JL-SiNWFET based biosensors can be used for effective diagnosis of biomarker diseases such as Ebola virus[30], gamma-Aminobutyric acid (GABA)[31] and breast cancer[32][33][26].

ACKNOWLEDGMENT

This study was supported by Debre Tabor University, Ministry of Science and Higher Education of Ethiopia and Delhi Technological University, India.

REFERENCES

[1] P. S. Raja, R. J. Daniel, and N. Bino, "Silicon Nanowire Based MOSFET Parameter Analysis," *OSR J. Electron. Commun. Eng.*, vol. 3, no. 6, pp. 35–43, 2012.

[2] M. Ieong, B. Doris, J. Kedzierski, K. Rim, and M. Yang, "Silicon device scaling to the sub-10-nm regime," *Science (80-.).*, vol. 306, no. 5704, pp. 2057–2060, 2004.

[3] Z. Huang, A. Gao, S. Chen, Y. Wang, and T. Li, "Highly Sensitive Juntionless Nanowire Transistor Biosensor in Detecting Breast Tumor Marker," *Proc. IEEE Sensors*, vol. 2018–Octob, pp. 1–4, 2018.

[4] D. Passeri, A. Morozzi, K. Kanxheri, and A. Scorzoni, "Numerical simulation of ISFET structures for biosensing devices with TCAD tools," *Biomed. Eng. Online*, vol. 14, no. 2, p. S3, 2015.

[5] M. I. B. Chowdhury, M. J. Islam, M. J. Islam, M. M. Hasan, and S. U. Farwah, "Silvaco TCAD based Analysis of Cylindrical Gate -All-Around FET Having Indium Arsenide as channel and Aluminium Oxide as Gate Dielectrics," *J. Nanotechnol. its Appl. Eng.*, vol. 1, no. 1, pp. 1–12, 2016.

[6] M. W. Rony, P. Bhowmik, H. R. Myler, and P. Mondol, "Short channel effects suppression in a dual-gate gate-all-around Si nanowire junctionless nMOSFET," *Proc. 9th Int. Conf. Electr. Comput. Eng. ICECE 2016*, pp. 538–541, 2017.

[7] N. Trivedi, M. Kumar, S. Haldar, S. S. Deswal, M. Gupta, and R. S. Gupta, "Analytical modeling simulation and characterization of short channel Junctionless Accumulation Mode Surrounding Gate (JLAMSG) MOSFET for improved analog/RF performance," *Superlattices Microstruct.*, vol. 100, pp. 1263–1275, 2016.

[8] A. B. Siddik, N. M. M. Hossain, S. Quader, and M. I. B. Chowdhury, "Silicon on metal technology merged with cylindrical gate all around FET for enhanced performance," *3rd Int. Conf. Electr. Inf. Commun. Technol. EICT 2017*, vol. 2018–Janua, no. December, pp. 1–4, 2018.

[9] A. Goel, S. Rewari, S. Verma, and R. S. Gupta, "Dielectric Modulated Triple Metal Gate All Around MOSFET (TMGAA)for DNA Bio-Molecule Detection," *2018 IEEE Electron Devices Kolkata Conf.*, vol. 1, pp. 337–340, 2019.

[10] N. M. M. Hossain, S. Quader, A. B. Siddik, and M. I. B. Chowdhury, "TCAD based performance analysis of junctionless cylindrical double gate all around FET up to 5nm technology node," *20th Int. Conf. Comput. Inf. Technol. ICCIT 2017*, pp. 1–4, 2017.

[11] Y. Pratap, M. Kumar, S. Kabra, S. Haldar, R. S. Gupta, and M. Gupta, "Analytical modeling of gate-all-around junctionless transistor based biosensors for detection of neutral biomolecule species," *J. Comput. Electron.*, vol. 17, no. 1, pp. 288–296, 2018.

[12] Y. Wang and G. Li, "Simulation of a silicon nanowire FET biosensor for detecting biotin/streptavidin binding," *2010 10th IEEE Conf. Nanotechnology, NANO 2010*, pp. 1036–1039, 2010.

[13] Y. He, C. Fan, and S. T. Lee, "Silicon nanostructures for bioapplications," *Nano Today*, vol. 5, no. 4, pp. 282–295, 2010.

[14] R. Gautam, M. Saxena, S. Member, A. In, and G. A. A. Mosfet, "Numerical Model of Gate-All-Around MOSFET With Vacuum Gate Dielectric for Biomolecule Detection," *IEEE ELECTRON DEVICE Lett. VOL. 33, NO. 12, DECEMBER 2012*, vol. 33, no. 12, pp. 1756–1758, 2012.

[15] W. Deng, X. Ma, and J. Huang, "Surface potential calculation and drain current model for junctionless double-gate polysilicon TFTs," *AIP Adv.*, vol. 4, no. 8, pp. 1–12, 2014.

[16] T. Yu, R. Wang, and R. Huang, "Simulation of Line-Edge Roughness Effects in Silicon Nanowire MOSFETs," pp. 187–190, 2010.

[17] I. Khan, O. Morshed, and S. M. Mominuzzaman, "Transport characterization of silicon nanowire field effect transistors with strain and scattering," *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON*, vol. 2017–Decem, pp. 1666–1669, 2017.

[18] G. Jayakumar, K. Garidis, P. E. Hellström, and M. Östling, "Fabrication and characterization of silicon nanowires using STL for biosensing applications," *ULIS 2014 - 2014 15th Int. Conf. Ultim. Integr. Silicon*, no. C, pp. 109–112, 2014.

[19] N. Thapa and L. Maurya, "Performance advancement of High-K dielectric MOSFET," *Int. J. Innov. Adv. Comput. Sci.*, vol. 3, no. 3, pp. 98–103, 2014.

[20] S. S. Saib, S. Yadav, J. Rahul, A. Srivastava, and B. Raj, "Impact of HfO 2 in Graded Channel Dual Insulator Double Gate MOSFET," *J. Comput. Theor. Nanosci.*, vol. 12, no. 6, pp. 10–13, 2015.

[21] G. Wadhwa, "Label Free Detection of Biomolecules Using Charge-Plasma- Based Gate Underlap Dielectric Modulated Junctionless TFET," *J. Electron. Mater.*, 2018.

[22] D. S. Software, *Atlas User ' s Manual*, no. 408. 2016.

[23] S. K. Mohapatra, K. P. Pradhan, and P. K. Sahu, "Influence of High-k Gate Dielectric on Nanoscale DG-MOSFET," *Int. J. Adv. Sci. Technol.*, vol. 65, pp. 19–26, 2014.

[24] T. A. Faculty, H. K. Kim, and I. P. Fulfillment, "WORK FUNCTION TUNING OF ELECTRODE MATERIALS WITH SMALL MOLECULE SURFACE MODIFIERS," *Japanese J. Appl. Phys. 56, 124001*, no. August, 2017.

[25] S. Kumar, K. K. Paliwal, and S. Mahajan, "EFFECT OF REDUCTION IN GATE OXIDE 100NM MOSFET," *Int. J. Latest Res. Sci. Technol.*, vol. 7, no. 1, pp. 5–8, 2018.

[26] F. Puppo *et al.*, "SiNW-FET in-Air Biosensors for High Sensitive and Specific Detection in Breast Tumor Extract," *IEEE Sens. J.*, vol. 16, no. 10, pp. 3374–3381, 2016.

[27] S. L. Tripathi, R. Mishra, and R. A. Mishra, "Multi-Gate Mosfet Structures With High-K Dielectric Materials," *ournal of Electron Devices*, vol. 16, pp. 1388–1394, 2012.

[28] M. F. Transistor and Y. Taur, "ULSI Scaling Toward 10nm Gate-lengths: Challenges and Opportu- nities," 2019.

[29] E. Rahman, A. Shadman, and Q. D. M. Khosru, "Effect of biomolecule position and fi ll in factor on sensitivity of a Dielectric Modulated Double Gate Junctionless MOSFET biosensor," *Sens. Bio-Sensing Res.*, vol. 13, pp. 49–54, 2017.

[30] Y. Chen *et al.*, "Field-Effect Transistor Biosensor for Rapid Detection of Ebola Antigen," *Sci. Rep.*, vol. 7, no. 1, pp. 4–11, 2017.

[31] J.-H. Lee, E.-J. Chae, S. Park, and J.-W. Choi, "Label-free detection of γ-aminobutyric acid based on silicon nanowire biosensor," *Nano Converg.*, vol. 6, no. 1, 2019.

[32] P. Ambhorkar *et al.*, "Nanowire-based biosensors: From growth to applications," *Micromachines*, vol. 9, no. 12, pp. 1–19, 2018.

[33] P. Mohanty *et al.*, "Field Effect Transistor Nanosensor for Breast Cancer Diagnostics," 2012.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:35:54 UTC from IEEE Xplore. Restrictions apply.

2020 IEEE 20th International Conference on Nanotechnology (IEEE-NANO)
July 29-31 2020, Virtual Conference

Temperature investigation of a Novel 3nm TF-Bulk FinFET for Improved Performance

Mridul Prakash Kashyap and Rishu Chaujar*, Senior *Member, IEEE*

***Abstract*— In this investigation, a novel *Truncated Fin* (TF) Junctionless (JL) bulk FinFET is compared over it's conventional structure. This paper is the analysis of high-performance temperature characteristics of a (JL) FinFET at the scale of gate length 3 nm investigated on *Silvaco TCAD EDA Tools*. This (TF) FinFET is the truncation of the fin at it's lower end (near it's interface with substrate) to have good impact. Moreover, the material and width of substrate are decided such that there will be minimum leakage of drain current through it. The procurement of this work is reduction of SCEs marked as *S.S=66.5 mV/dec* and *DIBL=2.5 mV/V* at room temperature in case of TF-FinFET. The enhancement in device performance can be inspected by some more SCEs parameters, which actually vary when channel length get comparable to the depletion width of Source and drain junctions, such as drain current at OFF state (which found out to be decreased by *81.5 times* in TF-FinFET at 300K), substrate current (which stood lowest at 500K and highest at 200K), switching ratio denoted by "I_{ON}/I_{OFF}" (grown by *120 times* in case of TF-FinFET at 200K) and the threshold voltage "V_{th}" of the simulated devices (which was found to be 0.256 V for TF-FinFET and 0.148 V for conventional FinFET at 200K). These parameters are best observed between 200K to 300K.**

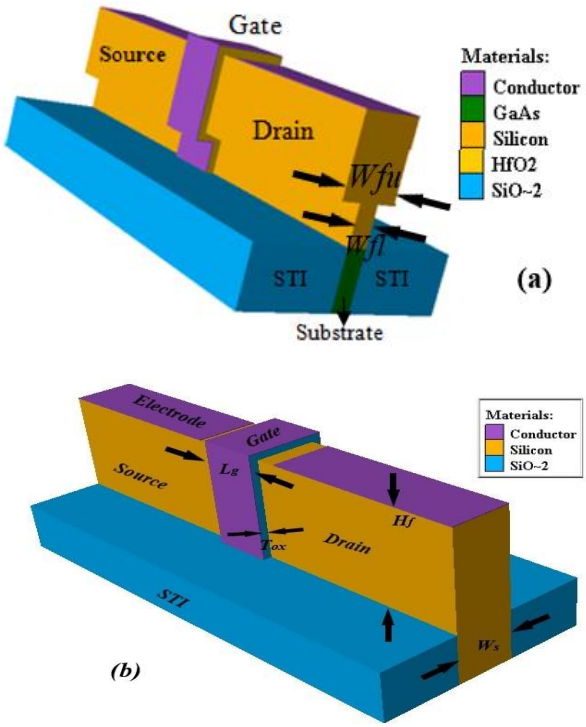
I. Introduction

During the last two decades, the FETs have been systematically scaled down in dimensions to achieve increased circuit density and high performance in a given chip area [1-3]. The idea of reducing the device dimensions while still managing the same current-voltage behavior as of comparatively larger device. Various studies have already been performed and explored intensively to minimize the impact of short channel effects (SCEs) such as Drain Induced Barrier Lowering (DIBL), Subthreshold swing (SS), low threshold voltage (V_{th}) [4-5]. In case of TF-FinFET, the I_{OFF} is lying in the order of 10^{-10} A/um and hence has the less discharging ability, which is good for high performance. In this TF-FinFET structure, silicon dioxide (SiO_2) is replaced with high-k Hafnium dioxide (HfO_2) as gate oxide. By changing the material (i.e., from Si to GaAs) and dimensions of substrate of this structure [6-7], the leakage of drain current comes out to be less because, the Energy band gap in GaAs is larger (E_g= 1.41eV) as compared to that of Silicon (E_g= 1.12eV).

Charges require more energy in GaAs (Gallium Arsenide) to jump from its E_v (valence band) to E_c (conduction band) as compared to Si (Silicon) [8]. This ultimately decreased I_{OFF} and hence improved the ratio of I_{ON}/I_{OFF}. The demand of special equipment is always being in demand, which can efficiently operate in the harsh conditions in the world (i.e., places with below 0° C and above 48° C). However, the device worked really well between 200K to 300K. This is evident to the utility of the device at room temperature as well as around 200K. Moreover, some characteristics [9-10] like Output conductance (g_{ds}), Early voltage (V_{ea}) and Substrate current at ON-state stood favorable at 500K. This ultimately showed the compatibility of TF-FinFET at such high temperature too.

II. Device Structure And Simulation Methodology

Fig.1 shows the proposed (i.e., TF-FinFET) and conventional structure respectively. Here the style of doping concentration in both the devices is kept same absolutely as mentioned later. Mainly the changes are made in material and structure of the proposed device. Due to this structural change, the device looks like somehow similar to conventional as well as GAA FET.



*Resrach supported by Delhi Technological University (DTU).
Dr. Rishu Chaujar* is with Micro research Lab, Department of Applied Physics, Delhi Technological University, New Delhi, 110042, India (*corresponding author* phone: +919811385965; e-mail: Chaujar.rishu@dtu.ac.in).

Mridul Prakash Kashyap is with Delhi Technological University, New Delhi, 110042, India. He is now with the Department of Applied Physics, Delhi Technological University, New Delhi, 110042 India (e-mail: mridul.prakash2012@gmail.com).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:16:55 UTC from IEEE Xplore. Restrictions apply.

Fig.1 (a) 3-D view of TF FinFET. (b) 3-D view of Conventional FinFET. Temperature variation of both is mentioned in later section.

In TF-FinFET, the fin is divided into upper fin (i.e., upper part of 4 nm height) and lower fin (i.e., lower truncated part of 2 nm height). The lower fin is thinner than upper fin by 2 nm, which ultimately optimized the gate control over channel region. As this device is simulated at the gate length of 3 nm, so other structural parameters are kept in accordance with it. Fin height is kept at 6 nm, lower fin width at 1nm and its upper width at 3 nm, workfunction of gate metal at 4.55 eV and gate oxide thickness at 0.4 nm. Doping concentration of the simulated n-type device is taken as $1.0x10^{20}$ (cm^{-3}) for source, $1.0x10^{18}$ (cm^{-3}) for channel, $1.0x10^{19}$ (cm^{-3}) for drain and $1.0x10^{15}$ (cm^{-3}) for substrate [11-12].

The oxide thickness is kept same at 0.4 nm in both the cases, but HfO2 (having k=25 approx.) is given preference over SiO2 (having k=3.9) in case of the TF-FinFET. This replacement enhanced the gate controllability (i.e., by giving higher driving current), however this replacement would surely have some trade-offs, such as time taken by the capacitors to get charge and discharge would increase for sure due to the *increment in gate capacitance of this device* [13]. Both Devices have same device length of 25 nm, same source/drain length of 11 nm, same device height of 9 nm and same gate length (Lg) of 3 nm. The value of electrode metal workfunction is taken as 4.55 eV for source, drain and gate metal electrode. The following physical models used under the drift-diffusion formulation are modification of local density approximation for quantum correction, such as conc. Dep. Mobility and Recombination models, Fermi-Dirac statistic (which describes the distribution of particles over energy states in a system consisting of many identical particles that obeys Pauli exc. principle), hydrodynamic energy transport model (which adds continuity equations for the carrier temp., treat mobilities and impact ionization coefficients as function of temp. rather than function of local electric fields) and Band to Band tunnelling model [14] (which shows the carrier movement through the bands which shouldn't able to move, when assumed classically).

TABLE I. STRUCTURAL PARAMETETS FOR SIMULATION

Parameters	Dimensions of Simulated Devices	
	TF-FinFET (n-type)	Conv. FinFET (n-type)
Gate length (L_g)	3nm	3nm
Fin height (H_f)	6nm	6nm
Lower fin width (W_{fl})	1nm	3nm
Upper fin width (W_{fu})	3nm	3nm
Effective oxide thickness (T_{ox})	0.4nm	0.4nm
Gate Electrode Thickness ($T_{elec.}$)	0.05nm	0.05nm
Gate metal workfunction	4.55	4.55
Contact resistivity	417nΩ.m	417nΩ.m
Substrate width (W_s)	1nm	3nm

III. SIMULATED RESULTS AND DISCUSSION

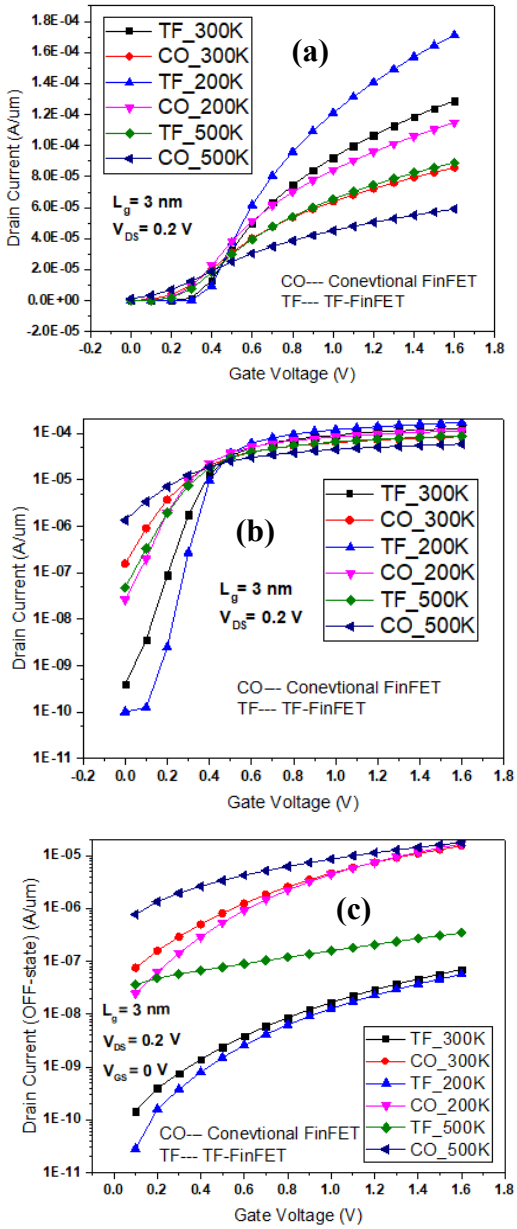


Fig.2.a) Variation of Drain current w.r.t V_G at V_D=0.2V. b) Variation of Drain current (log.) w.r.t V_G at V_D=0.2V. c) Variation of Drain current (OFF-state) w.r.t V_G at V_D=0.2V.

Fig.2 shows the variation of drain current at ON and OFF state at three different temperature (i.e., 200K, 300K and 500K) for both conventional and TF-FinFET devices. Here, one can easily depict the domination of TF-FinFET over the conventional FinFET, which is most favorable (i.e., max.) at 200K and 300K [15]. Moreover, the leakage current (i.e., I_D at OFF-state) also showed the influence of TF-FinFET at 200K and 300K over the conventional device. Conventional device at 500K showed the poorest result among all three cases. This ultimately evident the good existence of TF-FinFET at such a low temperature. This ultimately evident the robustness of TF-FinFET at these temperatures.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:16:55 UTC from IEEE Xplore. Restrictions apply.

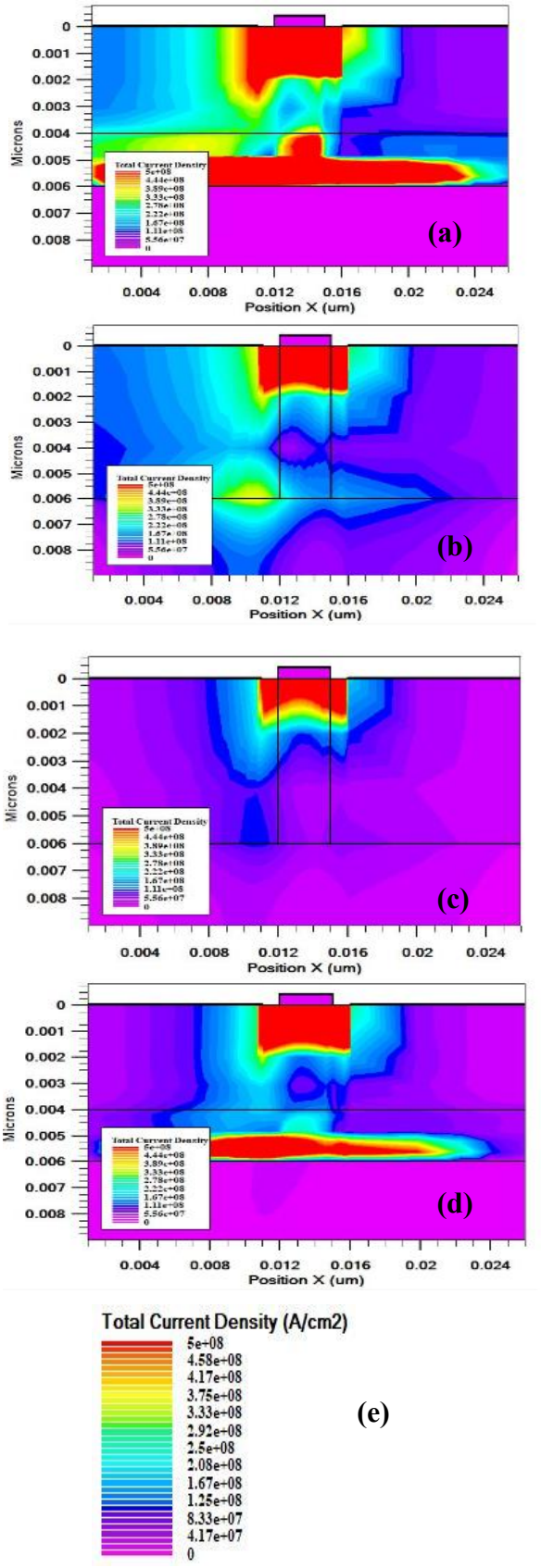


Fig.3 (a) Variance of current density in TF-FinFET at 200K. (b) Variance of current density in conventional FinFET at 200K. (c) Variance of current density in conventional FinFET at 500K. (d) Variance of current density in TF-FinFET at 500K. (e) Values of current density for fair comparison in all cases.

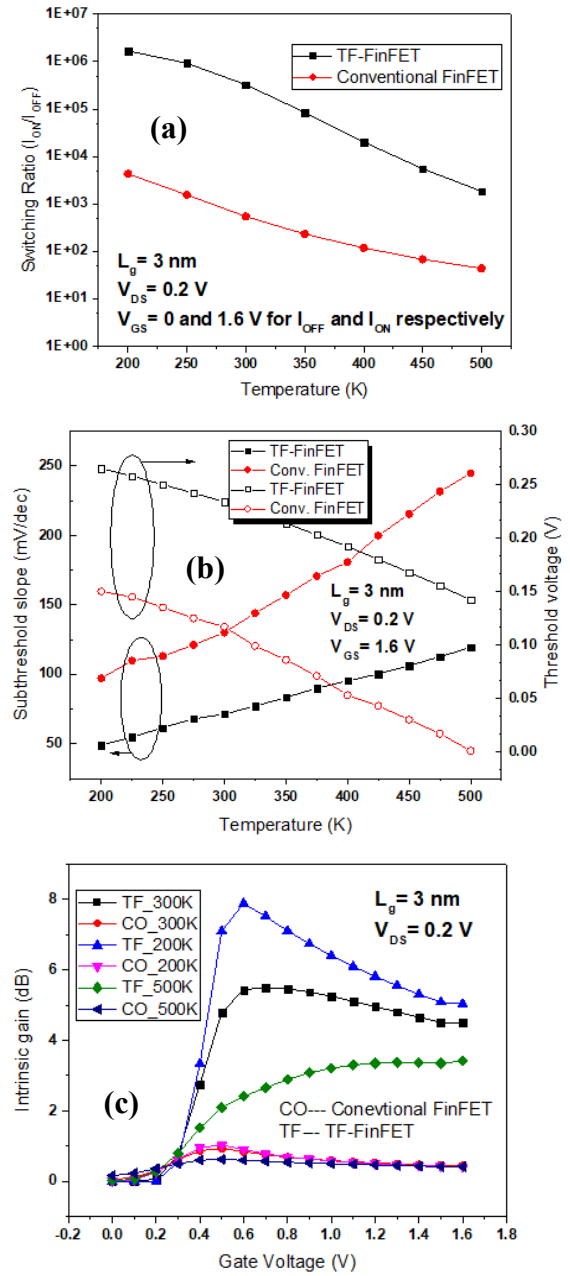


Fig.4 (a) Variation of switching ratio in both FinFETs w.r.t temperature. (b) Variance of S.S and V_{th} in both FinFETs at different temperatures. (c) Variation of Intrinsic gain in both the FinFET at w.r.t temperature.

Fig.3 shows the variation of total current density in both the devices at 200K and 500K. The devices at 200K have higher amount of current at their ON-state (i.e., at V_G=1.6 V) as compared to devices at 500K. Moreover, the leakage current in TF-FinFET at ON-state (i.e., substrate current) found very less in comparison to that of conventional FinFET at both the temperatures. This due to the TF-mechanism of TF-FinFET, as its truncation in the fin and substrate material (GaAs) helped the current to flow from drain to source with min. leakage [16-17].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:16:55 UTC from IEEE Xplore. Restrictions apply.

Fig.4 (a) and (b) shows the performance enhancement of TF-FinFET over conventional FinFET. The variation of switching ratio (I_{ON}/I_{OFF}) is far better for TF-FinFET over this wide range of temperature. Moreover, the Subthreshold swing (S.S) and V_{th} variation conclude the domination of TF-FinFET again. As, TF-FinFET has S.S<50 mV/dec at 200K, which is due to the fact that TF-mechanism has good control over the channel. Plot **(c),** ultimately conclude the feasibility of TF-FinFET between 200K to 300K. This is due to desired variation of transconductance (g_m) and Output conductance (g_{ds}) over this variety of temperature [18-19].

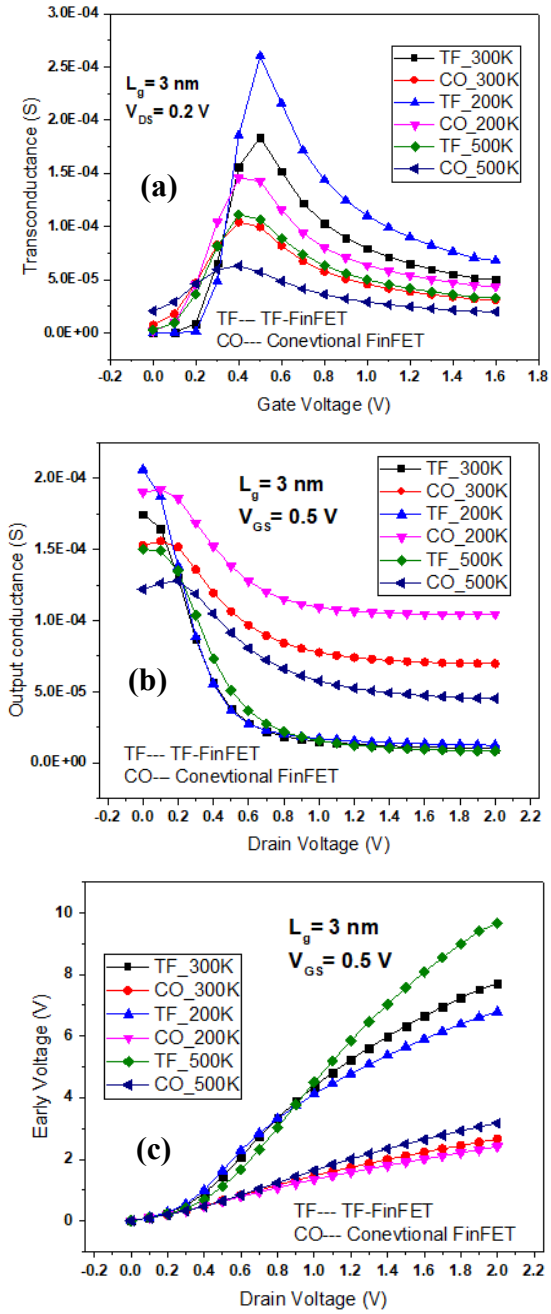


Fig.5 (a) Variation of Transconductance in both FinFETs w.r.t temperature. (b) Variance of Output conductance in both FinFETs at different temperatures. (c) Variation of Early voltage in both the FinFET w.r.t temperature.

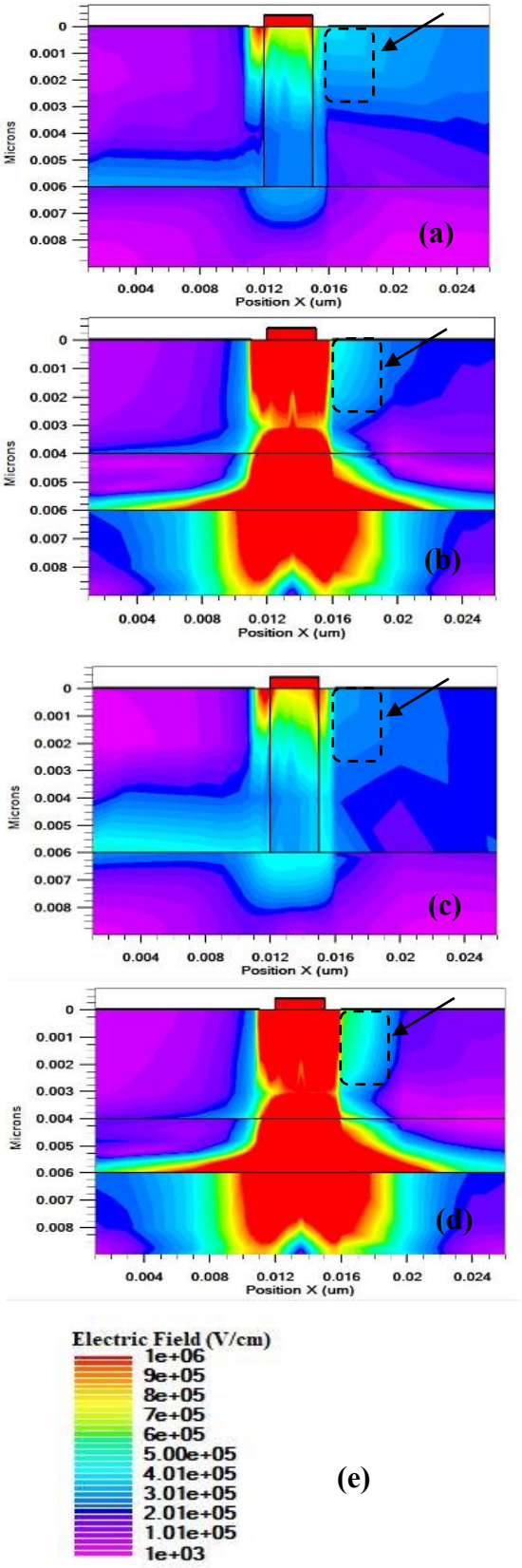


Fig.6 (a) Variance of electric field in Conventional FinFET at 200K. (b) Variance of electric field in TF-FinFET at 200K. (c) Variance of electric field in Conventional FinFET at 500K. (d) Variance of electric field in TF-FinFET at 500K. (e) Values of electric field for fair comparison in all cases.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:16:55 UTC from IEEE Xplore. Restrictions apply.

Fig.5 (a) and (b) evident the performance enhancement of TF-FinFET between the temperature 200K to 300K. As, we have large incremental change in drain current in TF-FinFET at 200K and 300K among all cases. So, due to which, one can easily depict the variation of Transconductance in both the devices at different temperatures. As far as Output conductance is concerned, the variation is same for all the cases of TF-FinFET. This is because, there would be minimum difference in output resistance of TF-FinFET over temperature variation [20]. Plot **(c),** depicts the early voltage for the devices for various temperatures. As analyzed, the TF-FinFET has more early voltage over conventional FinFET, which ultimately indicates that it has more amplifying capacity to the input signal [21].

Fig.6 evident the superiority of TF-FinFET over conventional FinFET in terms of *hot-carrier effect* near channel-drain interface. So less is the electric field at channel-drain interface, less will be the hot-carrier effect (i.e., leakage) from channel region. Here, TF-FinFET at 200 K has least hot carrier effect and hence has least leakage. But conventional FinFET at 500 K has minimum hot carrier effect and hence has minimum leakage as compared to TF-FinFET at 500 K [22].

As far as the ***manufacturing limitations*** are concerned, they could be overcome by improving the degree of precision in process of PR and lithography. As these are the explicit properties of manufacturing process, which are getting sophisticated w.r.t to time and whenever the precision will get introduce then, our design would ultimately become feasible on its own.

IV. Conclusion

In this investigation, temperature analysis at the gate length of 3nm scale at bulk FinFET is performed. The proposed (i.e., TF-FinFET) structure bear all the advantages of high-performance temperature results. The TF-FinFET worked really well between the temperature 200K to 300K. Decreased I_{OFF} in case of TF-FinFET (ranges in 10^{-11} at 200K) over the conventional FinFET (ranges in 10^{-8} at 200K) and switching ratio (grown by *120 times* in case of TF-FinFET) are the major achievements of this work. The premier results are analyzed in case of TF-FinFET between 200k to 300K, which ultimately showed the good performance of TF-FinFET at room temperature as well as at such a cold temperature (i.e., far below then 0° C). Simulated results showed enhanced performance of TF-FinFET in terms of System-on-chip (SOC) applications at temperature between 200K to 300K.

References

[1] N. Boukortt, S. Patanè, and G. Crupi, "3D Investigation of 8-nm Tapered n-FinFET Model," *Silicon*, pp. 19–20, 2019, DOI: 10.1007/s12633-019-00253-y.

[2] Narain Arora**.** Mosfet Modelling for VLSI Simulation., *Mosfet Modelling for VLSI Simulation.* International Series on Advances in Solid State Electronics and Technology,2007, pp. 90-93.

[3] A. Kumar, M. M. Tripathi, and R. Chaujar, "Investigation of parasitic capacitances of In2O5Sn gate electrode recessed channel MOSFET for ULSI switching applications," *Microsyst. Technol.*, vol. 23, no. 12, pp. 5867–5874, 2017, DOI: 10.1007/s00542-017-3348-2.

[4] A. Kumar, N. Gupta, and R. Chaujar, "Analysis of novel transparent gate recessed channel (TGRC) MOSFET for improved analog behaviour," *Microsyst. Technol.*, vol. 22, no. 11, pp. 2665–2671, 2016, DOI:10.1007/s00542-015-2554-z.

[5] A. Kumar, M. M. Tripathi, and R. Chaujar, "Reliability Issues of In2O5Sn Gate Electrode Recessed Channel MOSFET: Impact of Interface Trap Charges and Temperature," *IEEE Trans. Electron Devices*, vol. 65, no. 3, pp. 860–866, 2018, DOI:10.1109/TED.2018.2793853.

[6] E. Yu, K. Heo, and S. Cho, "Characterization and Optimization of Inverted-T FinFET under Nanoscale Dimensions," *IEEE Trans. Electron Devices*, vol. 65, no. 8, pp. 3521–3527, 2018, DOI:10.1109/TED.2018.2846478.

[7] A. B. Sachid, C. R. Manoj, D. K. Sharma, V. Ramgopal Rao, "Gate Fringe-Induced Barrier Lowering in Underlap FinFET Structures and Its Optimization," in IEEE Electron Device Letters, Vol. 29, No. 1, pp. 128-130, Jan. 2008, DOI: 10.1109/LED.2007.911974

[8] M. H. Han, C. Y. Chang, H. Bin Chen, J. J. Wu, Y. C. Cheng, and Y. C. Wu, "Performance comparison between bulk and SOI junctionless transistors," *IEEE Electron Device Lett.*, vol. 34, no. 2, pp. 169–171, 2013, DOI:10.1109/LED.2012.2231395.

[9] W. T. Huang and Y. Li, "Electrical characteristic fluctuation of 16-nm-gate trapezoidal bulk FinFET devices with fixed top-fin width induced by random discrete dopants," *Nanoscale Res. Lett.*, vol. 10, no. 1, pp. 1–8, 2015, DOI:10.1186/s11671-015-0739-0.

[10] Zheng, Peng & Liao, Yuehua & Damrongplasit, Nattapol & Chiang, Meng-Hsueh & King Liu, Tsu-Jae. (2014). "Variation-aware comparative study of 10-nm GAA versus FinFET 6-T SRAM performance and yield*". IEEE Transactions on Electron Devices*, vol. 61, no.12, pp. 3949-3954, 2014, 10.1109/TED.2014.2360351.

[11] D. Chatterjee and A. Kottantharayil, "A CMOS Compatible Bulk FinFET-Based Ultra Low Energy Leaky Integrate and Fire Neuron for Spiking Neural Networks," *IEEE Electron Device Lett.*, vol. 40, no. 8, pp. 1301–1304, 2019, DOI:10.1109/LED.2019.2924259.

[12] A. Kranti and G. A. Armstrong, "Design and optimization of FinFETs for ultra-low-voltage analog applications," *IEEE Trans. Electron Devices*, vol. 54, no. 12, pp. 3308–3316, 2007, DOI: 10.1109/TED.2007.908596.

[13] C. Liu, F. Zheng, Y. Sun, X. Li, and Y. Shi, "Novel tri-independent-gate FinFET for multi-current modes control," *Superlattices Microstruct.*, vol. 109, pp. 374–381, 2017, DOI:10.1016/j.spmi.2017.05.025.

[14] J. Madan and R. Chaujar, "Numerical Simulation of N+ Source Pocket PIN-GAA-Tunnel FET: Impact of Interface Trap Charges and Temperature," *IEEE Trans. Electron Devices*, vol. 64, no. 4, pp. 1482–1488, 2017, DOI: 10.1109/TED.2017.2670603.

[15] S. M. Nawaz, S. Dutta, A. Chattopadhyay, and A. Mallik, "Comparison of random dopant and gate-metal workfunction variability between junctionless and conventional FinFETs," *IEEE Electron Device Lett.*, vol. 35, no. 6, pp. 663–665, 2014, DOI:10.1109/LED.2014.2313916.

[16] S. Khandelwal, J. P. Duarte, Y. S. Chauhan, and C. Hu, "Modeling 20-nm Germanium FinFET with the industry standard FinFET model," *IEEE Electron Device Lett.*, vol. 35, no. 7, pp. 711–713, 2014, DOI: 10.1109/LED.2014.2323956.

[17] Kedzierski, Jakub & Ieong, Man & Nowak, E.J. & Kanarsky, Thomas & Zhang, Ying & Roy, Ronnen & Boyd, Diane & Fried, David & Wong., "Extension and source/drain design for high-performance FinFET devices," *IEEE Trans. Electron Devices*, vol. 50, no. 4, pp. 952–958, 2003, DOI:10.1109/TED.2003.811412.

[18] S. Gupta and A. Nandi, "Effect of air spacer on analog performance of underlap tri-gate FinFET," *Superlattices Microstruct.*, vol. 109, pp. 693–701, 2017, DOI:10.1016/j.spmi.2017.05.053.

[19] P. H. Su and Y. Li, "Analysis of In-Line Process Parameters of the Unity Gain Frequency of HKMG Bulk FinFET Devices," *IEEE Electron Device Lett.*, vol. 39, no. 3, pp. 335–338, 2018, DOI: 10.1109/LED.2018.2791436.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:16:55 UTC from IEEE Xplore. Restrictions apply.

[20] V. Thirunavukkarasu, Y. R. Jhan, Y. B. Liu, and Y. C. Wu, "Performance of Inversion, Accumulation, and Junctionless Mode n-Type and p-Type Bulk Silicon FinFETs with 3-nm Gate Length," *IEEE Electron Device Lett.*, vol. 36, no. 7, pp. 645–647, 2015,DOI:10.1109/LED.2015.2433303.
[21] ITRS 2.0 Publication, "International Technology Roadmap for Semiconductors, 2015," Available online at http://www.itrs2.net/.
[22] PH Drive, S. Clara, Atlas User's Manual, Device Structure Editing and Simulation Software; 2017. Pp. 1-1715. Available: www.silvaco.com

.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 08,2020 at 05:16:55 UTC from IEEE Xplore. Restrictions apply.

The Effect of Gate Stack and High-κ Spacer on Device Performance of a Junctionless GAA FinFET

Bhavya Kumar
Department of Applied Physics
Delhi Technological University
Delhi, India
bhavyakumar_phd2k18@dtu.ac.in

Ajay Kumar
ECE Department
Jaypee Institute of Information Technology
Noida, India
ajaydce@ieee.org

Rishu Chaujar
Department of Applied Physics
Delhi Technological University
Delhi, India
chaujar.rishu@dtu.ac.in

***Abstract* – In this work, we investigated the effect of the gate stack and high-κ gate spacers on the digital and analog performance of a lightly doped n-type Si channel Junctionless Rectangular Gate All Around (JL-Re-GAA) FinFET. Different digital and analog parameters, for instance, drain current (I_d), leakage current (I_{off}), switching ratio (I_{on}/I_{off}), subthreshold swing (SS), transconductance (g_m), output conductance (g_d), transconductance generation factor (TGF), intrinsic gain (A_v), early voltage (V_{EA}) have been analyzed. From the simulated results obtained, we have found that the use of gate stack and high-κ gate spacers remarkably improves the digital and analog figures of merits (FOMs) of the device. Thus, the JL-Re-GAA FinFET structure with high-k gate spacers and gate stack can be considered as a suitable candidate in digital and analog circuit applications.**

Keywords – Analog Figure of Merits (FOMs), FinFET, Gate Stack (GS), High-κ Gate Spacers, Junctionless Rectangular Gate All Around (JL-Re-GAA).

I. Introduction

As the size of the device decreases gradually, the performance of the device gets affected and unwanted effects like a subthreshold swing (SS), DIBL is introduced known as short channel effects (SCEs) [1]. To alleviate the SCEs, several multi-gate (MG) devices have been put forward and extensively researched. In 1999, professor Chenming Hu proposed a "3D" transistor named FinFET, which emerged as one of the most promising devices due to its (1) enhanced drive current per unit area, (2) higher switching (I_{on}/I_{off}) ratio (3) reduced SCEs due to superb electrostatic control by the multiple gates over the channel. However, the difficulty arises in the fabrication process as the fabrication of sharp junction at the nanoscale is a little tedious task, and to control the doping also requires extra effort, and these things might increase the cost of the device [2]. The solution to the above problems was provided by Colinge in the form of JLFET (Junctionless Field Effect Transistor) [3]. In JL transistors, doping species (either n-type or p-type) is the same in all the regions [4]. With the increase in no. of gates over the channel, electrostatic control of the devices increases resulting in higher I_{on} current, thus Gate All Around (GAA) structure was introduced. As per the 2015 ITRS guideline, the GAA structure is also known as "ultimate structure" [5,6].

Further, the GAA structure can be divided into two categories, Cylindrical Gate All Around (Cy-GAA) and Rectangular Gate All Around (Re-GAA) [7]. So, combining all the references, we are proposing Junctionless Rectangular Gate All Around (JL-Re-GAA) FinFET. When the gate oxide thickness (t_{ox}) decreases and reaches beyond 2 nm, gate direct tunneling current increases leading to an exponential rise in the off-state leakage current which degrades the device performance [8]. The high-κ dielectric material is used to overcome this problem. However, the direct deposition of high-κ dielectrics on Si substrate remains a challenging filed. An alternative is to use Gate Stack (GS) architecture comprising of a thin interfacial SiO_2 layer as passivation between the Si substrate and high-κ dielectrics by maintaining the effective oxide thickness (EOT) constant [9]. The high-κ gate spacers are also used to enhance the overall device performance.

In this work, extensive simulations have been performed to analyse the several important digital and analog parameters of the three different devices, namely, GAA FinFET, GAA FinFET with spacer, and GS-GAA FinFET with spacer. Starting with the introduction, section II narrates the device structure and simulation methodology used. Section III explains the digital and analog performance parameters of the device with concluding remarks in section IV authenticating the originality of the paper.

II. Device Structure and Simulation Methodology

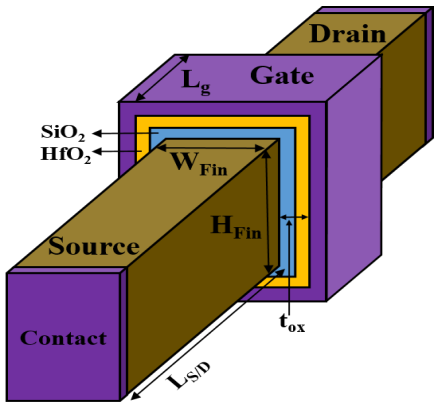


Fig. 1(a) Systematic 3-D structure of JL-Re-GAA FinFET with gate stack.

978-1-7281-1933-5/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:30:51 UTC from IEEE Xplore. Restrictions apply.

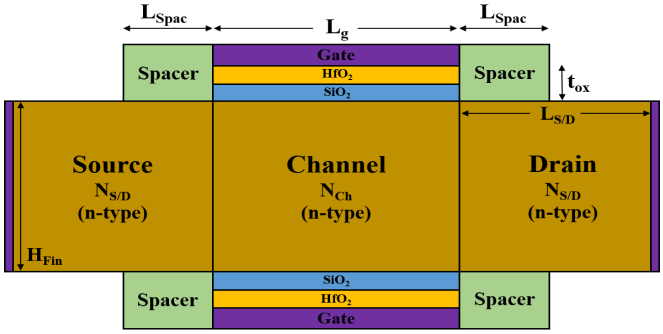


Fig. 1(b) 2-D cross-sectional view of JL-Re-GAA FinFET with gate stack and spacer.

Fig. 1(a) represents the proposed 3D view of JL-Re-GAA FinFET with gate stack, whereas Fig. 1(b) shows the 2D cross-sectional view of JL-Re-GAA FinFET with gate stack and spacer. Table I summarizes the device dimension of each structure. The gate length of the device (L_g) is fixed at 22 nm and the length of S/D regions ($L_{S/D}$) is fixed at 11 nm for all the device structures [7]. Oxide thickness (t_{ox}) is kept constant at 1 nm all around the fin. SiO_2 is used as the gate oxide in the first two configurations, whereas in gate stack configuration SiO_2 and HfO_2 are amalgamated in equal proportion. The fin is made up of silicon material. The height of the fin (H_{Fin}) is kept 4 times greater than the width of the fin (W_{Fin}) because the high Fin Aspect Ratio (H_{Fin}/W_{Fin}) achieves low leakage current (I_{off}), high I_{on} current, reduces SCEs. Thus, a high Fin Aspect Ratio (H_{Fin}/W_{Fin}) provides better digital and analog performance [2]. Donor doping is used to make the device n-type and since it is a junctionless device all the regions i.e., source, drain, and channel are uniformly doped with n-type doping. The doping concentration in the channel region (N_{Ch}) is 5×10^{16} cm^{-3}, whereas the doping concentration in the source and drain regions ($N_{S/D}$) is fixed at 10^{19} cm^{-3}. Source and drain regions are heavily doped in comparison to the channel to reduce the parasitic capacitance [2]. HfO_2 ($\kappa = 25$) is used as the high-κ gate dielectric spacer. Nowadays, metal gates are preferred over polysilicon gates because polysilicon shows a poly-depletion effect, thereby affecting the EOT of the device. The work function (ϕ_m) is 4.8 eV for each configuration.

All the simulations are carried out by the SILVACO ATLAS 3D simulator. During the entire simulations, drain voltage (V_{ds}) is kept at 0.5V, gate-source voltage (V_{gs}) is changed from 0-1V and the temperature has been set at 300K. In device simulation, various models are included for different purposes like the CVT model is incorporated to grab the field-dependent mobility and for recombination Shockley-Read-Hall (SRH) model is employed to imitate the existence of trap charges at the interface. To increase the precision in the results, the Fermi-Dirac Statistics model is used and to incorporate the energy transport effects, the Hydrodynamic model is included [10]. Further, Gummel and Newton's methods are applied to attain the solution.

TABLE I
DIFFERENT DEVICE PARAMETERS USED FOR SIMULATION

Parameter	GAA FinFET	GAA FinFET with spacer	GS-GAA FinFET with spacer
L_g (nm)	22	22	22
t_{ox} (nm)	1	1	1 (0.5 SiO_2 + 0.5 HfO_2)
H_{Fin} (nm)	40	40	40
W_{Fin} (nm)	10	10	10
$L_{S/D}$ (nm)	11	11	11
L_{Spac} (nm)	-	5	5
$N_{S/D}$ (cm^{-3})	$1x10^{19}$	$1x10^{19}$	$1x10^{19}$
N_{Ch} (cm^{-3})	$5x10^{16}$	$5x10^{16}$	$5x10^{16}$

III. RESULTS AND DISCUSSION

Fig. 2(a, b) represents drain current (I_d) variation with gate-source voltage (V_{gs}) at constant V_{ds} = 0.5V. For comparative analysis, I_d is calculated for Tri-Gate (TG) FinFET and is compared with other structures. The dimensions are the same for the Tri-Gate (TG) FinFET. A higher I_{on} current and lower I_{off} current is obtained for the GS-GAA FinFET with spacer structure when compared to their counterparts. The reason is that the use of high-κ gate spacers or high dielectric materials enhances the strength of fringing electric fields, which reduces the potential barrier in the OFF-state, resulting in increased ON-state current and decreased OFF-state current [11,12].

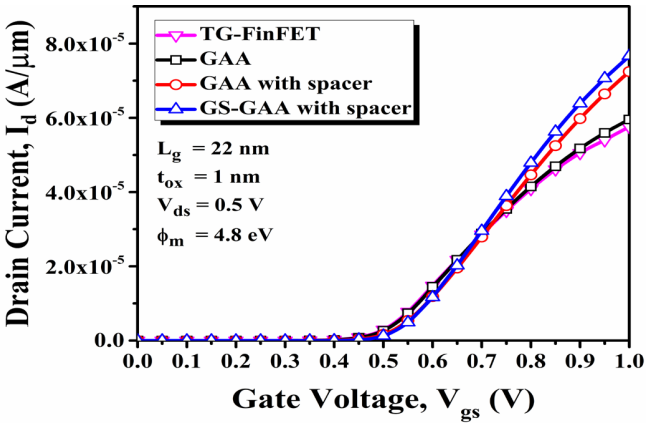


Fig. 2(a) Variation of I_d-V_{gs} in linear scale for different structures.

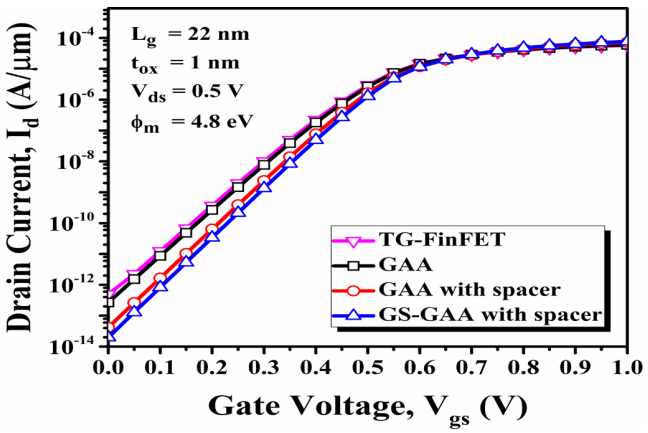


Fig. 2(b) Variation of I_d-V_{gs} in log scale for different structures.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:30:51 UTC from IEEE Xplore. Restrictions apply.

Fig. 3 and Fig. 4 represent the switching (I_{on}/I_{off}) ratio and threshold voltage (V_{th}) respectively, for each device structure. Better the I_{on}/I_{off} ratio, better the switching speed of the device. The GS-GAA FinFET with spacer structure has the highest I_{on}/I_{off} ratio and threshold voltage (V_{th}) when compared with other structures. This is because the strength of fringing electric fields increases with the high-κ spacer and in the OFF-state, silicon gets depleted far away from the gate edges, resulting in the improved subthreshold characteristics of the device [11].

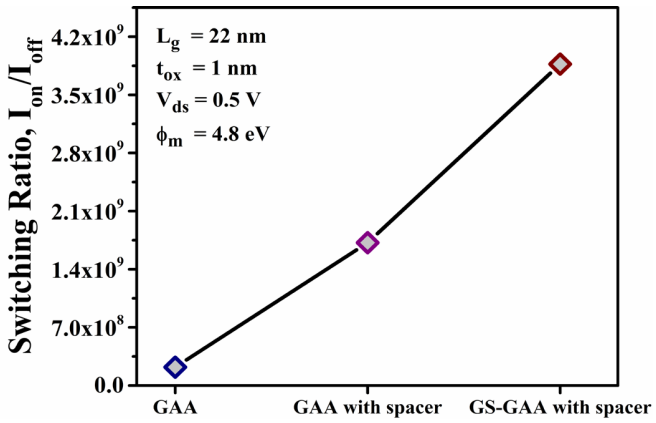


Fig. 3 I_{on}/I_{off} ratio comparison for different structures.

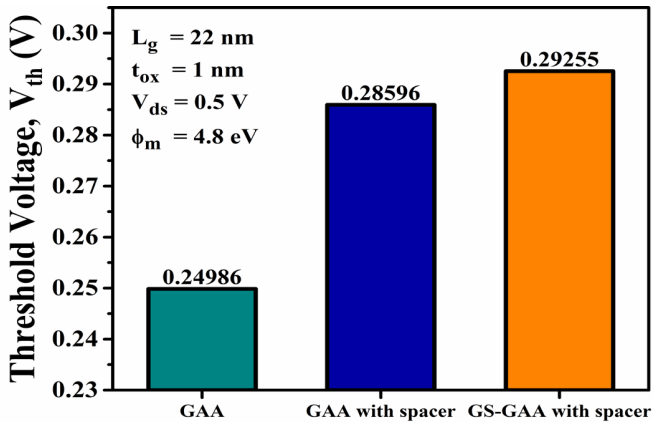


Fig. 4 Threshold Voltage (V_{th}) comparison for different structures.

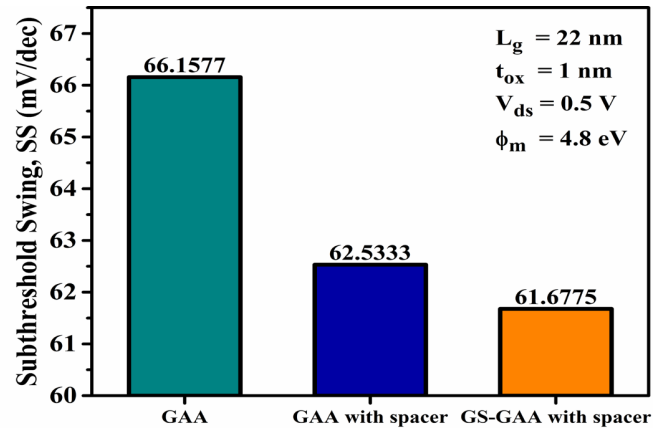


Fig. 5 Subthreshold Swing (SS) comparison for different structures.

The subthreshold swing (SS) is the gate-source voltage needed to alter the I_d by one order of magnitude per decade. In Fig. 5, SS is plotted for all the device structures. SS is closest to its ideal value 60 mV/dec for GS-GAA FinFET with spacer structure because as the dielectric constant increases, gate capacitance increases which improves the controllability of gate over the channel resulting in improved SS. Thus, the use of high-k spacers and gate stack concept overall improves the digital performance of the device in terms of I_{on}, I_{off}, I_{on}/I_{off} ratio, SS, V_{th}.

From the analog applications point of view, analog FOMs namely transconductance (g_m), transconductance generation factor (TGF), output conductance (g_d), intrinsic gain (A_v), early voltage (V_{EA}) are analyzed and calculated using Eq. (1) to Eq. (5) [13].

$$g_m = \frac{\partial I_d}{\partial V_{gs}} \tag{1}$$

$$TGF = \frac{g_m}{I_d} \tag{2}$$

$$g_d = \frac{\partial I_d}{\partial V_{ds}} \tag{3}$$

$$A_v = \frac{g_m}{g_d} \tag{4}$$

$$V_{EA} = \frac{I_d}{g_d} \tag{5}$$

The gate overdrive voltage (V_{gt}) is evaluated by the difference of gate-source voltage and threshold voltage i.e., ($V_{gt} = V_{gs} - V_{th}$). In amplifier circuits, the region of operation is decided by V_{gt}, as a result, all the analog analysis has been done against V_{gt} [13]. Transconductance as defined in Eq. (1) is a parameter that relates the shift in the I_d to the shift in V_{gs}. In Fig. 6, we get the maximum peak of g_m for GS-GAA FinFET with spacer structure because high-k spacer increases the electron velocity and current density, thereby enhancing the electron mobility and consequently g_m increases.

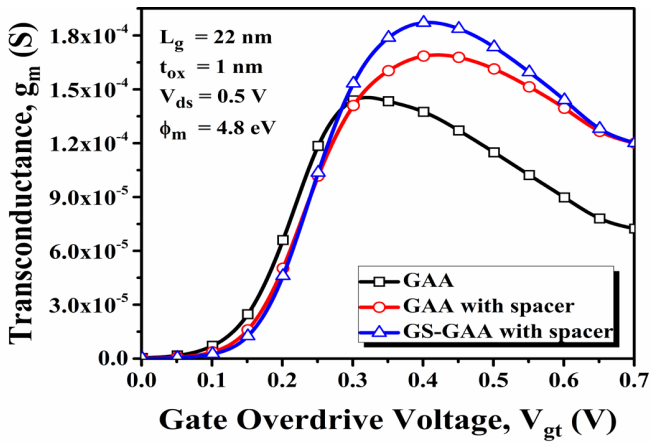


Fig. 6 Transconductance vs gate overdrive voltage for different structures.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:30:51 UTC from IEEE Xplore. Restrictions apply.

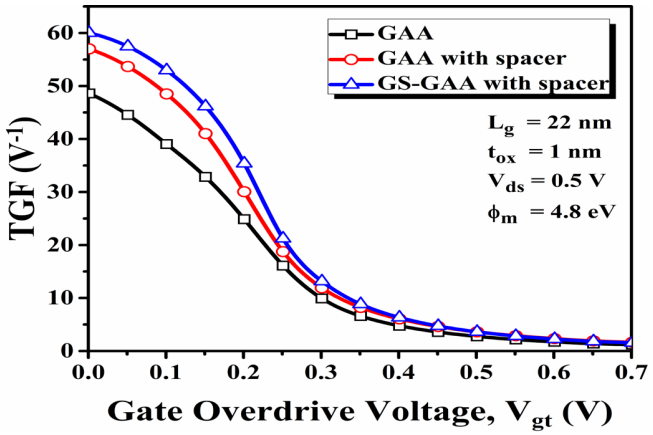


Fig. 7 Transconductance generation factor (TGF) variation as a function of gate overdrive voltage (V_{gt}) for different structures.

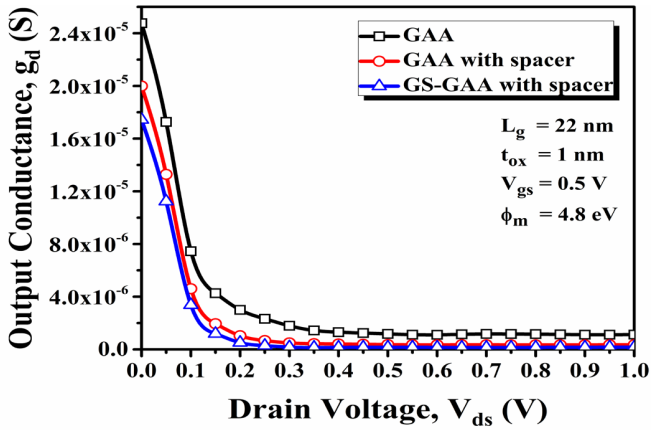


Fig. 8 Output conductance (g_d) vs drain voltage (V_{ds}) at constant V_{gs} = 0.5V for different structures.

TABLE II
VALUES OF DIGITAL AND ANALOG PARAMETERS FOR DIFFERENT STRUCTURES

Device Structure	I_{on} (A/ µm)	I_{off} (A/ µm)	V_{th} (V)	SS (mV/dec)	g_m (S)	TGF (V^{-1})	g_d (S)	A_v (dB)	V_{EA} (V)
GAA FinFET	5.95×10^{-5}	2.68×10^{-13}	0.2498	66.15	1.43×10^{-4}	48.60	2.47×10^{-5}	124.70	52.85
GAA FinFET with spacer	7.24×10^{-5}	4.19×10^{-14}	0.2859	62.53	1.68×10^{-4}	57.00	1.99×10^{-5}	475.43	205.06
GS-GAA FinFET with Spacer	7.67×10^{-5}	1.98×10^{-14}	0.2925	61.67	1.87×10^{-4}	60.02	1.74×10^{-5}	1303.92	448.89

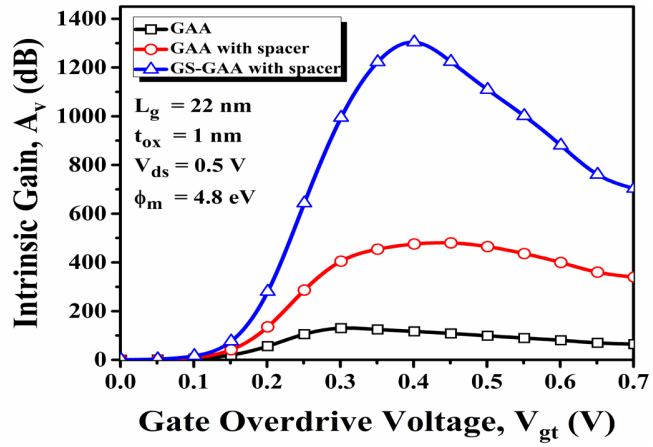


Fig. 9 Intrinsic gain (A_v) vs gate overdrive voltage (V_{gt}) for different configurations.

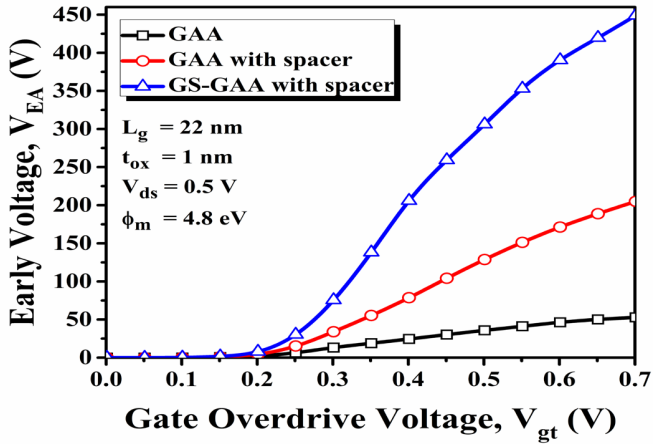


Fig. 10 Variation of early voltage (V_{EA}) as a function of gate overdrive voltage (V_{gt}) for different configurations.

The transconductance generation factor (TGF) is the amount of gain generated per unit power loss as expressed in Eq. (2). Higher the TGF better the device will work at the low power supply. From Fig. 7, it can be seen that TGF is maximum for GS-GAA FinFET with spacer structure in the subthreshold region because higher I_d corresponds to higher g_m and hence higher TGF. The Current driving ability of a device is determined by output conductance (g_d) as defined in Eq. (3). In Fig. 8, output conductance is plotted against V_{ds} at constant V_{gs} = 0.5V. Output conductance decreases as the V_{ds} increases and approaches zero as the device reaches in the saturation region.

Output conductance decreases due to the shift in the lateral electric field from the gate side to the drain side, which occurs due to the increase in a fringing electric field [14] and, as a result, g_d is lowest for GS-GAA FinFET with spacer structure as demonstrated in Table II. The intrinsic gain (A_v) of the device as calculated in Eq. (4) is the ratio of g_m and g_d. The variation of intrinsic gain against V_{gt} is displayed in Fig. 9 for each configuration. In Fig. 6 and Fig. 8, high g_m and low g_d are obtained for GS-GAA FinFET with spacer structure. Hence, the ratio gets much bigger and as a result, we get a maximum peak of A_v for GS-GAA FinFET gate stack with a spacer structure.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:30:51 UTC from IEEE Xplore. Restrictions apply.

The early voltage as defined in Eq. (5) is the ratio of drain current and output conductance. In Fig. 10, V_{EA} is plotted against V_{gt} for each configuration. Due to the reduction in SCEs, an improvement is observed in early voltage for the GS-GAA FinFET with spacer structure. In gate stack structure EOT of the device reduces and as a result, V_{EA} increases and drain current dependence on the drain voltage V_{ds} reduces. For better performance of the device, intrinsic gain (A_v) and early voltage (V_{EA}) must be very high.

IV. Conclusion

In this work, the effect of the gate stack and high-κ gate spacers on the digital and analog performance of the Junctionless Rectangular Gate All Around (JL-Re-GAA) FinFET has been investigated using the SILVACO ATLAS 3D Simulator. From the simulated results, it is clear that the GS-GAA FinFET with the spacer structure shows the most improved results. In GS-GAA FinFET with spacer structure, I_{on} current get increased by 28.9% and subthreshold swing (SS) decreased by around 6.77%, whereas switching ratio (I_{on}/I_{off}) increased by over one order of magnitude and I_{off} current decreased by around one order of magnitude when compared with the Gate All Around (GAA) FinFET without gate stack and spacer. It is also observed that analog FOMs show improved results for GS-GAA FinFET with spacer structure with g_m get enhanced by 30.76%, TGF by 23.49%, and g_d get decreased by 29.55%, whereas A_v and V_{EA} increased by almost ten times. Hence, the Junctionless Rectangular Gate All Around (JL-Re-GAA) FinFET with gate stack and spacer structure not only improves the reliability and performance parameters of the device discussed above but also provides better immunity to SCEs.

Acknowledgment

The authors are grateful to Microelectronics Research Laboratory, Deptt. of Applied Physics, and Delhi Technological University for supporting the work.

References

[1] A. Chhabra, A. Kumar, and R. Chaujar, "GaAs Junctionless FinFET using Si3N4 Spacer for High performance Analog Application," in *2018 International Conference On Advances in Communication and Computing Technology (ICACCT)*, 2018, pp. 483–486.

[2] K. Biswas and C. K. Sarkar, "Optimizing Fin Aspect Ratio of Junctionless Bulk FinFET for Application in Analog/RF Circuit," in *2018 IEEE Electron Devices Kolkata Conference (EDKCON)*, 2018, pp. 591–595.

[3] J. P. Colinge, C. W. Lee, A. Afzalian, N. D. Akhavan, R. Yan, I. Ferain, P. Razavi, B. O'Neill, A. Blake, M. White, A. M. Kelleher, B. McCarthy, and R. Murphy, "Nanowire transistors without junctions," *Nat. Nanotechnol.*, vol. 5, no. 3, pp. 225–229, 2010.

[4] K. Biswas, A. Sarkar, and C. K. Sarkar, "Spacer engineering for performance enhancement of junctionless accumulation-mode bulk FinFETs," *IET Circuits, Devices Syst.*, vol. 11, no. 1, pp. 80–88, 2016.

[5] *The International Technology Roadmap for Semiconductors 2.0: 2015*, Itrpv, 2015 [Online]. Available: http://www.itrs2.net/.

[6] Y. C. Huang, M. H. Chiang, S. J. Wang, and J. G. Fossum, "GAAFET Versus Pragmatic FinFET at the 5nm Si-based CMOS Technology Node," *IEEE J. Electron Devices Soc.*, vol. 5, no. 3, pp. 164–169, 2017.

[7] M. K. Rai, V. Narendar, and R. A. Mishra, "Significance of variation in various parameters on electrical characteristics of FinFET devices," *SCES 2014 Inspiring Eng. Syst. Glob. Sustain.*, vol. 1, pp. 1–6, 2014.

[8] H. Iwai and H. S. Momose, "Ultra-thin gate oxides - performance and reliability," *Tech. Dig. - Int. Electron Devices Meet.*, pp. 163–166, 1998.

[9] N. Gupta and R. Chaujar, "Optimization of high-k and gate metal workfunction for improved analog and intermodulation performance of Gate Stack (GS)-GEWE-SiNW MOSFET," *Superlattices Microstruct.*, vol. 97, pp. 630–641, 2016

[10] *ATLAS User's Manual,* SILVACO International, Santa Clara, CA, USA, 2016.

[11] R. K. Baruah and R. P. Paily, "A Dual-Material Gate Junctionless Transistor with high-k spacer for enhanced analog performance," *IEEE Trans. Electron Devices*, vol. 61, no. 1, pp. 123–128, 2014.

[12] I. V. Singh and M. S. Alam, "Analog/RF performance evaluation of nanoscale non-overlap SOI MOSFETs with high-k stack on spacer," *CODEC 2012 - 5th Int. Conf. Comput. Devices Commun.*, vol. 3, pp. 1–4, 2012.

[13] S. K. Mohapatra, K. P. Pradhan, P. K. Sahu, and M. R. Kumar, "The performance measure of GS-DG MOSFET: An impact of metal gate work function," *Adv. Nat. Sci. Nanosci. Nanotechnol.*, vol. 5, no. 2, 2014.

[14] A. Nandi, A. K. Saxena, and S. Dasgupta, "Impact of dual-k spacer on analog performance of underlap FinFET," *Microelectronics J.*, vol. 43, no. 11, pp. 883–887, 2012.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2020 at 05:30:51 UTC from IEEE Xplore. Restrictions apply.

2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)
Amity University, Noida, India. June 4-5, 2020

Twitter Sentiment Analysis Using Lexical or Rule Based Approach: A Case Study

Sheresh Zahoor
Electronics and Communication Department
Delhi Technological University
Delhi, India
sahrishzahoor@gmail.com

Rajesh Rohilla
Electronics and Communication Department
Delhi Technological University
Delhi, India
rajesh@dce.ac.in

Abstract— **Opinion analysis or sentiment analysis is one of the most effective techniques these days in order to determine the sentiments or emotions of people regarding any event. This technique has come to the fore because of extensive use of social media platforms like Facebook, Twitter etc. by people to express their emotions regarding any event that has occurred or any event that is most likely to happen, be that the release of a movie or a political rally that is about to take place. People make sure to express their sentiments. Sentiment analysis proves very beneficial for any company selling a product to know how their product was received by people or by any political party to determine how people are reacting towards their running candidate. In order to analyze these sentiments two approaches of machine learning can be used – unsupervised or supervised. In this paper we have worked with Unsupervised Approach which is a rule based or lexical approach and can be done using the pre-built open source libraries like TextBlob, VADER.**

Keywords—**Twitter Sentiment analysis; Unsupervised approach; TextBlob; VADER; #HaryanaAssemblyPolls; #MLKhattar; #TheSkyIsPink; #UNGA**

I. INTRODUCTION

Opinions have always been of utmost importance in fields like politics, art, entertainment, fashion etc. where every aspect works on opinions and sentiments of people related to these fields or events that occur every day throughout the world. Before the advent of the social networking sites like Twitter, Facebook, Instagram etc., people relied on newspapers, television sets, magazines for news or any big event occurring in the world, but now with social media people have all the information in their hands and are not dependent on a single source for their news or opinions. Social media has drastically changed the outlook of life of many people and has given them a platform to express their views related to any event or any product. Social media allows a person to express his like or dislike towards any product, person etc.

Researches indicate that using the social media sites is considered as the best way to grow a business in terms of money, time, effort and other resources [2]. However, the massive availability of opinions and the unstructured nature in which these opinions are available, makes it very difficult to extract any useful information out of them, so over the years it has become very important to devise and develop techniques that could ease extracting information out of the unstructured opinions that are present on any social networking site.

Sentiment analysis, also referred to as opinion mining is one of the most recent techniques that has been developed in order to extract information gathered from any social networking site. Sentiment analysis is an automated process of computationally identifying and classifying the opinions expressed in a piece of text to determine whether the sentiment of a person towards a particular event, product, person etc. is positive, negative or neutral. Sentiment analysis is considered the most advanced way of determining opinions as it makes use of AI to classify a piece of text into positive, negative or neutral.

Twitter is one of the fastest growing social media sites and has become very popular over the last few years as people openly express their opinions in the form of tweets with a maximum length of 140 characters. As of 2019, Twitter had around 330 million active users. Of these, more than 40 percent, or more specifically, 134 million, use the service on a daily basis [3]. This accounts for a large amount of unstructured data which can prove to be very useful to determine opinions or sentiments.

This is a case study of a number of events that have occurred throughout the world and involves determining how people perceived these events. It involves the implementation of Twitter sentiment analysis by utilizing APIs provided by Twitter itself. Tweets were collected related to events like Howdy Modi, a gathering that occurred in Houston; Haryana Assembly Elections; a movie release i.e. The sky is pink, Delhi odd-even traffic management and the international conference UN General Assembly. The data of all these events was scraped from twitter and was analyzed using an open source Python library called TextBlob and another in-built library NLTK-VADER.

II. LITERATURE REVIEW

Sentiment analysis refers to the study of text on various levels. Document level [9] This is the first level where objects are classified according to their attributes. (Turney, 2002; Pang and Lee, 2004), The second level is the sentence level wherein [11] the sentences are classified on the basis of whether they represent negative, positive or neutral sentiments (Hu and Liu, 2004; Kim and Hovy, 2004). The next level is the phrase level [12] wherein it is determined whether the expression is unbiassed or polar, followed by the removal of uncertainty.

Pak and Paroubek (2010) [5] classified tasks as subjective and objective. For subjective data, the information was gathered from the user tweets such as text or image or symbols. For objective information on the other hand, the information was gathered from authentic sources such as newspapers.

Information which is taken for analysis is simply a sample of flowing tweets collected by using queries. In the past years there have been numerous documents observing

978-1-7281-7016-9/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 21,2020 at 05:08:54 UTC from IEEE Xplore. Restrictions apply.

the Twitter sentiment and buzz [5], [12] (Jansen et al. 2009; Pak and Paroubek 2010; O'Connor et al. 2010; Tumasjan et al. 2010; Bifet and Frank 2010; Barbosa and Feng 2010; Davidov, Tsur, and Rappoport 2010). Barbosa and Feng (2010) [4] analyzed the sentimental classification on Twitter data by determining the polarity of tweets on the basis of symbols, retweets, emoticons and even syntax of the Tweets. Kamps et al. (2002) analyzed the data by using the lexical database such as WordNet which is the description of lexemes. This contains the emotional content of a word. The distance metric of words is used to determine semantic polarity of adjectives.

III. SENTIMENT ANALYSIS

Sentiment Analysis is one of the most recently developed techniques of determining opinions of people. It is a contextual mining of text which extracts information from a piece of text. Sentiment Analysis or "opinion mining" refers to the use of Natural Language Processing to determine the opinions and emotions of a piece of text. Sentiment Analysis is used to identify opinions, sentiments, attitudes, and states of mind of people towards a product or service and classify them as positive, negative and neutral. Sentiment analysis can be done using two Approaches-Unsupervised or Lexical or Rule based approach and Supervised or Machine Learning Based Approach.

A. Unsupervised or Lexical based Approach

In this method the text is represented as a bag of words and these words are then compared to the contents of dictionaries which contain words along with its associated semantic orientation. It is considered to be the unsupervised approach of sentiment analysis and is efficient in determining the polarity i.e. whether a given piece of text is positive, negative or neutral. The most commonly used library for sentiment analysis using lexical or rule based approach is TextBlob [4]. It is a python library for processing of textual data and simplest way of diving into common natural language processing tasks like sentiment analysis, classification. VADER is another such library; however, it has been specifically designed for analysis of text of social-media as the language of social-media is different from that of normal spoken or written language.

B. Supervised Approach

This approach is based on machine learning algorithms like Naïve Bayes, SVM etc. to determine the polarity of a piece of text.

In this research the lexical or rule based approach has been used for a number of events. For each of these events a minimum of 5000 tweets were scraped from the twitter, all this data is in the raw form i.e. it is unlabeled and there are no output and input parameters. Using the lexical approach, sentiments of all these tweets can be found out in a simple and efficient way. Two libraries i.e. TextBlob and VADER have been used for this purpose.

TextBlob- It is an open source Python library that is based on NLTK (Natural Language Tool Kit) and is used for analyzing a piece of text and assigning a polarity score to it. Polarity is a float value between [-1, 1] where 0 indicates a neutral sentiment, 1 indicates positive and -1 indicates negative sentiment. TextBlob analyses each word in a piece of text and assigns a semantic score to each word, then this score is weighted i.e. a weighted average is computed to assign a score to the complete sentence, the score is based on the polarity of each word in the sentence. TextBlob computes another important factor, Subjectivity- it is a float value having a range [0, 1] where 0 indicates that the sentence is objective which means that it is based on factual data whereas 1 indicates that the sentence is subjective which means that it is based on emotions, feelings, opinions, desires, and allegations of a person. It is an efficient way of analyzing raw tweets or any piece of text.

VADER-Valence Aware Dictionary and sentiment Reasoner is another python library that can be used for lexicon based sentiment analysis. However, the difference between VADER and TextBlob is that VADER was created mostly to analyze the social media texts i.e. it is able to comprehend and understand the social media slangs and language better than TextBlob and gives better result with social media texts than TextBlob. It is capable of analyzing repetitive vocabulary and the correlation coefficient shows that VADER ($r = 0.881$) performs as well as individual human raters ($r = 0.888$) at matching ground truth (aggregated group mean from 20 human raters for sentiment intensity of each tweet) [1]. Thus, VADER is considered to be highly efficient for social media text analysis. VADER breaks down the sentiment intensity score into a positive, negative and neutral component, these are further normalized to be within the range [-1, 1] as a compound score.

IV. METHODOLOGY

Lexical or rule-based approach is based on a number of steps that are very critical in extracting the sentiment or opinion from the raw tweets obtained from Twitter. These steps are:

- Data collection
- Data pre-processing
- Part of speech tagging (POS)
- Sentiment analysis using in-built dictionary

A. Data Collection

In order to be able to collect data from Twitter it is important to create an account on Twitter and then go to developer.twitter.com/en/apps to create an app that allows collection of tweets based on the query. However, one needs permission from Twitter to become a developer and after getting the permission, customer API keys are provided which are to be kept safe for further use while scraping the tweets. For one-time collection of tweets REST API is used and in this case study in order to collect the tweets based on different queries REST API has been used. Also, the data has been stored in the CSV files in this research. The CSV file will have the following columns:

- Created (date)
- Text (the actual tweet)
- Retweet
- Hashtag
- Followers

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 21,2020 at 05:08:54 UTC from IEEE Xplore. Restrictions apply.

- Friends

A. *Data Pre-Processing*

Before the actual sentiment analysis techniques are applied, a number of pre-processing steps are performed so as to make the tweet ready to be analyzed because the tweets obtained are in unstructured form and cannot be used as such without any processing. The typical pre-processing steps are as follows:

1) Tokenization or Bag-of-Words Creation: Tokenization involves separation of words from URLs, hashtags, at-mentions. Thus it is a method of breaking a string, text into a list of tokens. It is very important in the sentiment analysis as this gives an idea about the words that constitute a tweet. It is done in order to: Count the number of words in a piece of text and count the frequency of that word i.e. how many times that word was used or repeated in the sentence

2) N-grams Extraction: In this technique accompanying words are grouped together into phrases called n-grams. It has been observed that when words are grouped together, it improves the quality of sentiment analysis; however, there is no solution to the size of n-gram to be used.

3) Stemming and Lemmatization: In stemming words will be replaced by their stems or roots i.e. the words like write, writer and writing are mapped to a single word i.e. write and they are counted together as a result of which the dimensionality of the bag of words is reduced. In case of lemmatization, all the verbs are transformed to the infinite tense whereas all the nouns are transformed to the singular form.

4) Stop words removal: Stop words i.e. the prepositions, articles have a high frequency of occurrence but they do not affect the final sentiment of the text and thus should be removed from the text while pre-processing. For this purpose, the NLTK stop words in-built library can be used.

B. *Part-of-speech Tagging (POS)*

This is the process of automatically tagging each word in the text in terms of the part of speech it belongs to such as noun, pronoun, adverb etc.

C. *Sentiment analysis using built- in library*

In the lexical approach, sentiment analysis relies on an in-built library or dictionary of words with pre-calculated polarity. This method is considered as a part of Machine Learning i.e. unsupervised approach where the unlabeled data is analyzed to determine the polarity or sentiment of a tweet. In this research two in-built libraries have been used to analyze the sentiment of particular events and the results obtained are compared with each other. However, VADER is a library specifically designed for the social media texts or messages, so it has a much better hold at analyzing data from social media platform. The idea behind the lexicon approach is-

Constructing a lexicon of words with polarities

Forming a bag-of-words from the given text

Pre-processing the data like the removal of stop words etc.

Comparing each word with the lexicon and assigning a polarity to it

Calculating the whole sentiment score for the text by adding the sentiment score or polarity of each word in the text

V. CASES

A. *Haryana Assembly polls*

These assembly polls were held in Haryana, India in the month of October and the final turnout was 68.47% [6]. Tweets for this event were collected prior to the event and post the event to understand the overall sentiment of people towards the election and the leading party. Around 16,000 tweets with the hashtag #HaryanaAssemblyPolls were collected and analyzed using two in built libraries i.e. TextBlob and VADER as shown in Fig.1,2.

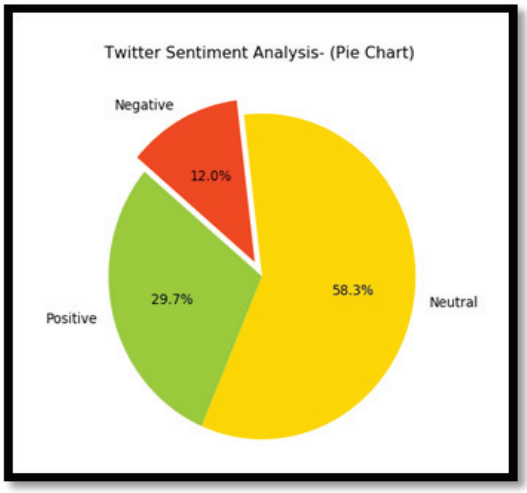


Fig. 1. Sentiment Analysis using TextBlob

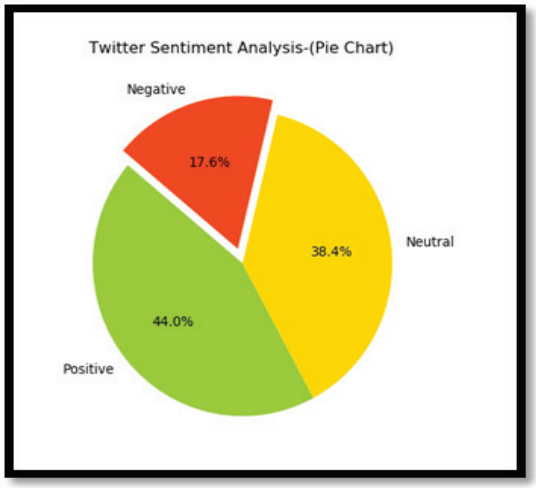


Fig. 2. Sentiment Analysis using VADER

It was noted as per the trending Hashtags that a single party was mentioned more than any other party i.e. BJP. So the tweets for that particular party and their candidate were scraped from Twitter to analyze the sentiment of people towards the BJP and their candidate ML Khattar. Around

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 21,2020 at 05:08:54 UTC from IEEE Xplore. Restrictions apply.

4000 tweets for BJP and ML Khattar were scraped and analyzed using TextBlob as well as VADER. It was found that the overall sentiment towards ML Khattar was positive.

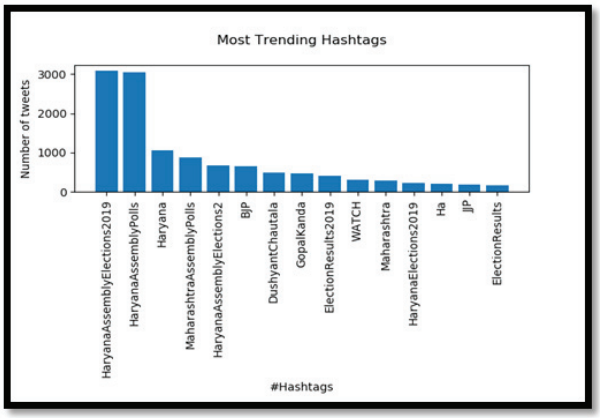


Fig. 3. Most Trending Hashtags

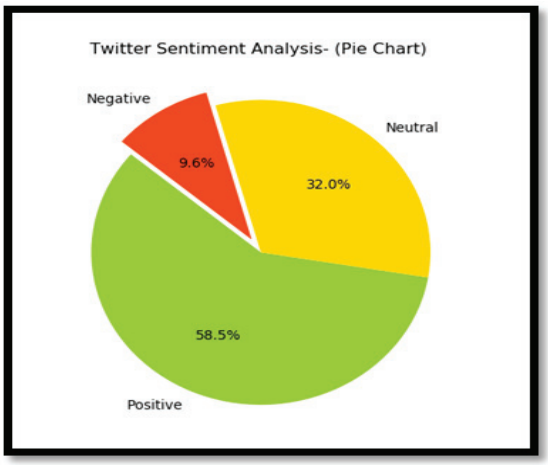


Fig. 4. Sentiment Analysis of ML Khattar using TextBlob

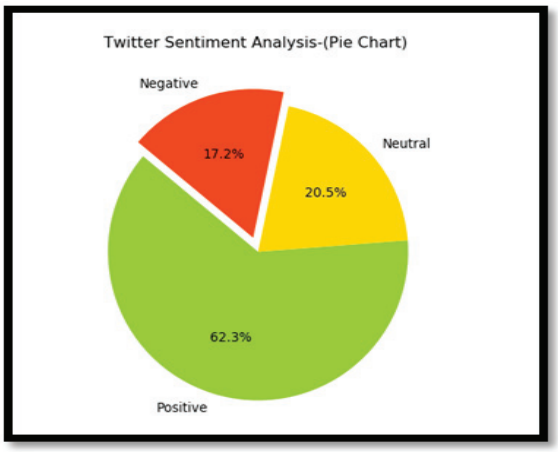


Fig. 5. Sentiment Analysis of ML Khattar using VADER

B. *The Sky Is Pink*

This movie starring Priyanka Chopra Jonas, Farhan Akhtar, Zaira Wasim and Rohit Suresh Saraf was released on 11 October. Around 8000 tweets for this particular movie were scraped using the hashtag #TheSkyIsPink. This movie gained popularity because it was considered to be a comeback movie of Priyanka and when hashtags related to this movie were scraped from Twitter, it was seen that #PriyankaChopra was leading.

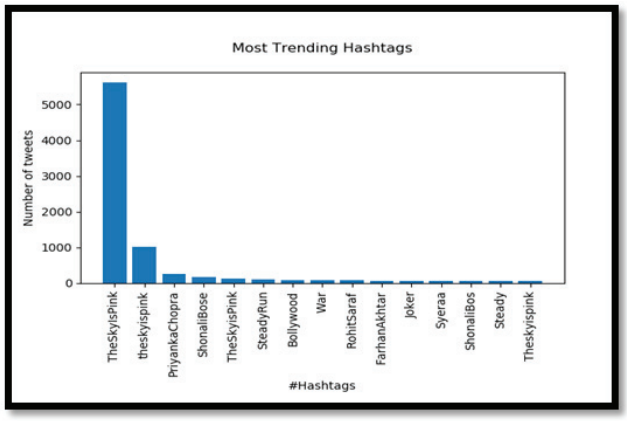


Fig. 6. Most Trending Hashtags

The tweets scraped were analyzed using TextBlob and VADER to determine the sentiment and the overall positive sentiment detected using TextBlob was 64.1% and using VADER was 62.8%. thus the movie was received positively by the audience and it could be seen that the most common hashtag amongst all was that of #PriyankaChopra.

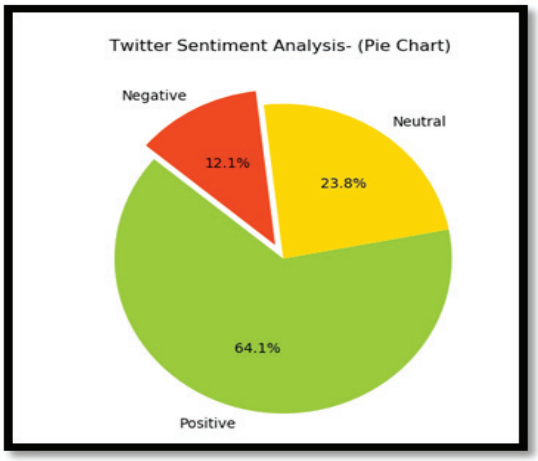


Fig. 7. Sentiment Analysis of The Sky Is Pink using TextBlob

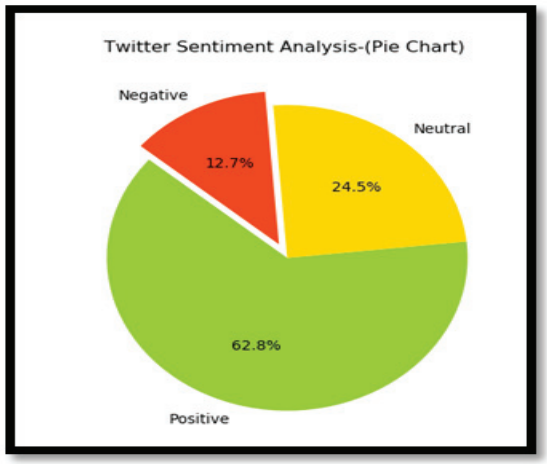


Fig. 8. Sentiment Analysis of The Sky Is Pink using VADER

C. *UNGA*

74th session of United Nations General Assembly started on 17th September 2019 and concluded on 24th September. Many political leaders delivered speeches addressing a number of issues faced by the world presently like the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 21,2020 at 05:08:54 UTC from IEEE Xplore. Restrictions apply.

economic slowdown, poverty, climate change, terrorism and many more. Many political issues were also brought to the table like that of Kashmir by the PM of Pakistan. Issues like sanitation and climate change were discussed by the PM of India. Many such issues of utmost importance were discussed by the world leaders. The trending hashtag was #UNGA. Around 10,000 tweets were collected and analyzed using the in-built libraries in order to determine the positive, negative or neutral sentiment of people towards this assembly.

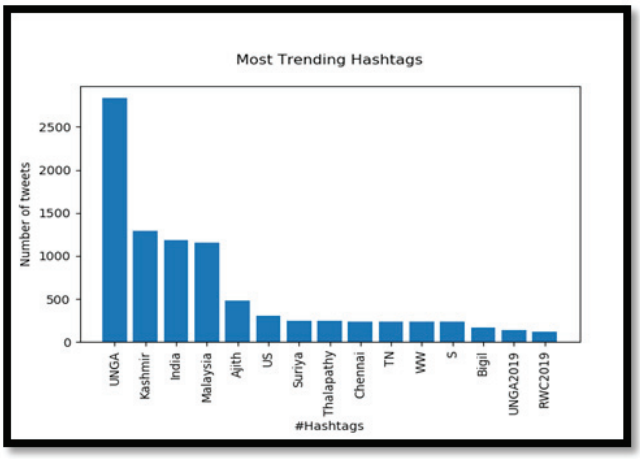


Fig. 9. Most Trending Hashtags

It was analyzed that using TextBlob the around 12.6% tweets were categorized as negative whereas using VADER around 40.1% tweets were categorized as negative. This case confirms the fact that the social media language somewhat affects the exact analysis, this is because of the use of abbreviations and slang in the social media text.

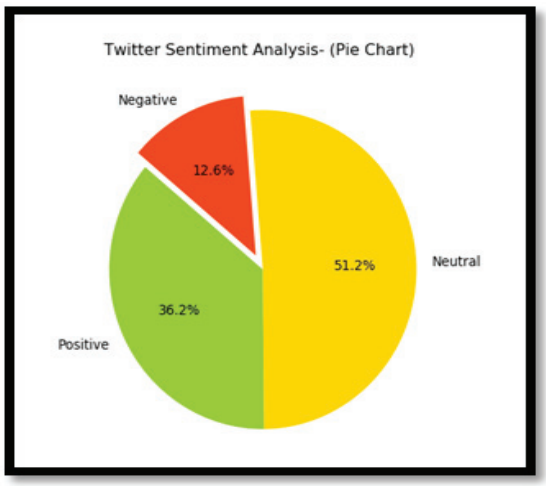


Fig. 10. Sentiment Analysis of UNGA using TextBlob

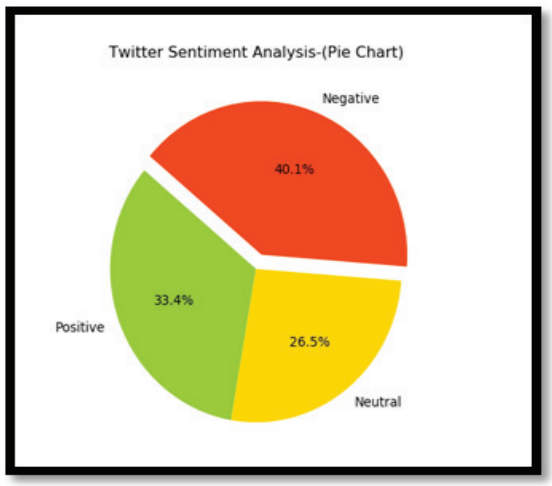


Fig. 11. Sentiment Analysis of UNGA using VADER

VI. RESULT

Results obtained have been tabulated in Table 1.

VII. CONCLUSIONS

It can be concluded that the results obtained using the lexical or rule-based approach (unsupervised learning) are not accurate and are subject to change because of the use of slangs in the social media text. Thus the raw data after getting classified into positive, negative or neutral using the manual or in-built libraries needs to be analyzed using a supervised machine learning algorithm in order to obtain the exact results and accuracy. Once the data is structured, it is divided into training and test set and then fed to the model for the classification. The supervised method is considered to be more specialized as compared to the unsupervised methods which make use of the in-built libraries. It is the supervised approach accompanied by unsupervised approach that needs to be developed further in order to obtain the exact results. Thus, Sentiment analysis has a vast scope of development in the future.

ACKNOWLEDGMENT

I would like to express my gratitude towards all the people who have contributed their precious time and effort to help me, without whom it would not have been possible for me to understand and complete the project.

I would like to thank Dr. Rajesh Rohilla, Department of Electronics and Communication Engineering, my guide for his support, motivation and encouragement throughout the period of this work.

TABLE I. SENTIMENT ANALYSIS

Sentiment	Haryana Assembly Polls		ML Khattar		The Sky is Pink		UNGA	
	TextBlob	*VADER*	*TextBlob*	*VADER*	*TextBlob*	*VADER*		
Positive	29.7%	44%	58.5%	62.3%	64.1%	62.8%	36.2%	33.4%
Negative	12.0%	17.6%	9.6%	17.2%	12.1%	12.7%	12.6%	40.1%
Neutral	58.3%	38.5%	32.%	20.5%	23.8%	24.5%	51.2%	26.5%

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 21,2020 at 05:08:54 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] C.J. Hutto Eric Gilbert, VADER: A Parsimonious Rule-based Model for Sentiment Analysis For Social Media Text

[2] Edosomwan, Simeon, et al. "The history of social media and its impact on business." Journal of Applied Management and entrepreneurship 16.3 (2011): 79-91.

[3] https://www.oberlo.in/blog/twitter-statistics

[4] https://textblob.readthedocs.io/en/dev/

[5] A. Pak and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining," in Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10) (N. C. C. Chair), K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, M. Rosner, and D. Tapias, eds.), (Valletta, Malta), European Language Resources Association (ELRA), may 2010

[6] http://results.eci.gov.in/ACOCT2019/partywiseresult-S07.htm

[7] "Twitter-sentiment-analysis-usingpython." Twitter Sentimental analysis for Realdonaldtrump" Devaki P, Ilakiya, J, Indumathi, R and Arul Priya, M

[8] Jansen,B.J.; Zhang,M.; Sobel,K.; and Chowdury,A. (2009), "Twitterpower: Tweets as electronic word of mouth", Journal of the American Society for Information Science and Technology 60(11):2169–2188..

[9] Pang, B., and Lee, L. (2008), "Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval" 2(1-2):1– 135.

[10] T. T. P. Souza, O. Kolchyna, P. C. Treleaven, and T. Aste, "Twitter sentiment analysis applied to finance: A case study in the retail industry," 2015

[11] M Hu and B Liu. (2004),"Mining and summarizing customer reviews. KDD".

[12] Wilson, T. Wiebe, J.; and Hoffmann, (P. 2009),"Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis. Computational Linguistics", 35(3):399–433.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 21,2020 at 05:08:54 UTC from IEEE Xplore. Restrictions apply.

Pramana – J. Phys. (2020) 94:129
https://doi.org/10.1007/s12043-020-01996-9
© Indian Academy of Sciences

Viscous Ricci dark energy model with matter creation: Exact solution and observational tests

AJAY KUMAR and C P SINGH*

Department of Applied Mathematics, Delhi Technological University, Delhi 110 042, India
*Corresponding author. E-mail: cpsingh@dce.ac.in

MS received 22 April 2020; accepted 18 June 2020

Abstract. In this paper, the dissipative mechanism (bulk viscosity and matter creation) is introduced to describe the effects of cosmic non-perfect fluid on the Ricci dark energy (RDE) model. We consider matter creation and bulk viscosity as two independent irreversible processes. Assuming suitable forms of the bulk viscous coefficient and matter creation rate, we find the exact solution of the field equations. We carry out fitting analysis on the cosmological parameters in the model by using Type Ia supernovae data, observational Hubble data and baryon acoustic oscillation (BAO) data with cosmic microwave background. We plot the trajectory of cosmological parameters with the best-fit values of model parameters and discuss all possible (deceleration, acceleration and their transitions) evolutions of the model. The current values of deceleration parameter and equation of state parameter are found to be $q_0 = -0.362$ and $\omega_{\text{eff}} = -0.575$, respectively. The age of the Universe is found to be $t_0 \simeq 13.397$ Gyr, which is very close to the ΛCDM model. We further discuss the geometrical diagnostic parameters such as statefinder and Om to distinguish the model with ΛCDM model. Finally, we discuss the behaviour of energy conditions for our model and find that the model satisfies the null energy condition (NEC), weak energy condition (WEC) and dominant energy condition (DEC) while it violates strong energy condition (SEC).

Keywords. Holographic dark energy cosmology; bulk viscosity; matter creation.

PACS Nos **95.36.+x; 98.80.−k**

1. Introduction

The study of accelerated expansion of the present Universe becomes one of the vital topics in cosmology during the last few years [1–3]. This fact is widely confirmed by the recent observational data. Many theories are considered to get a proper explanation of this phenomenon. These observations lead to a new kind of matter which violates the strong energy condition (SEC), i.e., $\rho + 3p < 0$. The energy density of matter responsible for such a condition to be satisfied at a certain stage of evolution of the Universe is referred to as dark energy (DE). Various theories have been proposed to explain DE problems [4–12]. The simplest DE model which is based on cosmological constant is dubbed ΛCDM model [13]. However, this model suffers from two problems called fine-tuning and coincidence problems.

Based on holographic principle [14], Li [15] proposed a new DE model called holographic dark energy (HDE) model in which the DE density was considered as $\rho_d = 3b^2 M_P^2 L^{-2}$, where b^2 is a dimensionless constant, $M_P^{-2} = G$ is the Planck mass and L denotes the infrared (IR) cut-off. It was found that the HDE with Hubble horizon as its IR cut-off does not give an accelerating Universe. Later on, he considered the future event horizon of the Universe as the infrared cut-off. This HDE model is able to explain the accelerated expansion of the Universe which fits the current observations. In the recent past, several works [16–21] have been carried out in HDE models to describe the accelerating behaviour of the Universe. Gao *et al* [22] suggested that the DE density might be inversely proportional to the Ricci scalar curvature, R, and they called this model the Ricci dark energy (RDE) model. In RDE model, the future event horizon has been replaced by the inverse of the Ricci scalar curvature. This model not only avoids the causality problem and is phenomenologically viable, but also solve the coincidence problem. They introduced HDE density proportional to the Ricci scalar, i.e., $\rho_d \propto R$, where $R = -6(2H^2 + \dot{H} +$

Published online: 20 August 2020

$\frac{k}{a^2}$). Thus, the energy density of the RDE model in a flat Friedmann–Robertson–Walker (FRW) Universe is given by

$$\rho_d = 3\alpha(\dot{H} + 2H^2), \tag{1}$$

where α is a dimensionless parameter, H is the Hubble parameter and $\dot{H}$ represents the derivative of the Hubble parameter with respect to the cosmic time t. This model fits with the observational data and well explain the coincidence problem also. In [23–26], researchers have shown that the RDE is suitable for describing the current acceleration of the Universe.

It is well known that the content of the Universe as a perfect fluid is assertive as it suggests no dissipation, which actually exists widely, and intuitively plays an important role in the evolution of the Universe. To be more realistic, the Universe is assumed to be filled with dissipative media. Therefore, cosmology based on imperfect fluid is proposed reasonably. Viscosity is a concept in fluid mechanics which is related to an exotic fluid with some thermodynamical features such as bulk and/or shear viscosities. In cosmology, shear viscosity is related to the anisotropy of the space–time whereas the bulk viscosity is usually related to the isotropic Universe. To reduce the equilibrium pressure in an expanding Universe, bulk viscosity can be useful. In cosmology, the viscosity concept was first discussed by Misner [27]. Several researchers [28–43] have discussed the effect of bulk viscosity to understand the DE phenomenon.

The study of matter creation in the relativistic cosmological models has drawn the attention of a number of researchers. In the framework of general relativity (GR), the adiabatic irreversible matter creation was first studied by Prigogine *et al* [44,45]. In their papers, they discussed that the second law of thermodynamics may be modified to accommodate flow of energy from gravitational field to the created matter field. This phenomenon of matter creation has been studied by many researchers in detail within the context of standard GR [46–54]. In the context of the recent acceleration, the concept of irreversible particle creation has been reconsidered due to its capability to produce an effective negative pressure. For more details, the reader is referred to refs [55–65]. Prigogine *et al* [44,45] considered the viscous and matter creation processes as two independent processes. Some researchers [35,66–68] have studied theses two dissipative processes by considering two independent phenomena.

In this paper, we study the effects of bulk viscosity and adiabatic matter creation in RDE model within the framework of FRW Universe. We discuss the evolution of the Universe by constraining the model parameters through combined observational data. We plot the trajectory of Hubble parameter and ΛCDM with error bar from Hubble data. We study the dynamical properties of our model by calculating the deceleration parameter, jerk, snap and lerk parameters analytically and geometrically. We further discuss the statefinder and Om diagnostics to discriminate RDE model with other standard DE models. We also discuss the behaviour of energy conditions for the RDE model.

This paper is organised as follows: In §2, we present the Einstein field equations with bulk viscosity and matter creation for RDE model in the framework of a flat FRW line element. In §3, we obtain exact solution to the field equations to obtain the Hubble parameter and the scale factor. We fit the model to the combined data set of Type Ia Supernovae, observational Hubble data and combined data of baryon acoustic oscillations and cosmic microwave background, and present the fitting results in §4. Many interesting issues including the age of the Universe are discussed in §5. We compare RDE model with other DE models by calculating statefinder parameters, Om diagnostics and cosmographic parameters in §6. In §7, we discuss the energy conditions for our model. Finally, we summarise our results in §8. It is to be noted that we have used matter creation and particle creation synonymously throughout the article.

2. The cosmological model

We consider a homogeneous and isotropic flat Friedmann–Robertson–Walker (FRW) line element with the units $8\pi G = 1$ and $c = 1$

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a^2(t)[\mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)], \tag{2}$$

where r, θ and ϕ are dimensionless co-moving coordinates and $a(t)$ is the scale factor. The Einstein's field equations (EFE) take the usual form

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = -T_{\mu\nu}, \tag{3}$$

where $T_{\mu\nu}$ is the energy–momentum tensor for the perfect fluid which modifies due to particle creation in adiabatic irreversible open thermodynamic systems and bulk viscous stress. We assume that the Universe is filled with pressureless dark matter and the energy of RDE for which the energy–momentum tensor is given by [68–70]

$$T_{\mu\nu} = (\rho_m + \rho_d + p_{\mathrm{eff}})u_\mu u_\nu - p_{\mathrm{eff}}g_{\mu\nu}. \tag{4}$$

Here, ρ_m and ρ_d denote the energy densities of dark matter and RDE, respectively, and p_{eff} is the effective pressure

$$p_{\mathrm{eff}} = p_d + p_c + p_v, \tag{5}$$

where p_d is the pressure of RDE, p_c is the pressure associated with the creation of particle out of the gravitational field [44,45] and p_v represents the viscous pressure which is assumed as $p_v = -3\zeta H$, where ζ is the bulk viscous coefficient [71]. The bulk viscous pressure, p_v, represents only a small correction to the thermodynamical pressure; it is a reasonable assumption that the inclusion of viscous term in the energy–momentum tensor does not fundamentally change the dynamics of the cosmic evolution. The pressure and density of RDE are related by the equation of state (EoS) $p_d = \omega_d \rho_d$, where ω_d is the EoS parameter of RDE.

In the background of metric (2) in the general theory of relativity, EFEs (3) yield the following two independent equations:

$$3H^2 = \rho_m + \rho_d, \tag{6}$$

$$2\dot{H} + 3H^2 = -p_{\text{eff}} = -(p_d + p_c - 3\zeta H), \tag{7}$$

where $H = \dot{a}/a$ is the Hubble parameter which measures the fractional rate of change of the scale factor $a(t)$ and the overhead dot indicates the derivative with respect to the cosmic time t.

In the presence of a gravitational particle source, the number of fluid particles is not conserved, i.e., $N^{\mu}_{;\mu} \neq 0$, where $N = nV$ is the total number of particles in a system. It means that N is variable, i.e., the system is open. Here, n is the particle number density and $V = a^3$ is the volume of the system. The particle number density flow $N^{\mu} = nu^{\mu}$ is assumed to satisfy the balance equation [68]

$$\dot{n} + 3Hn = n\Gamma, \tag{8}$$

where Γ denotes the particle creation rate of cold dark matter. Equation (8), when combined with the second law of thermodynamics, naturally leads to the appearance of a negative pressure directly associated with the rate Γ and the creation pressure p_c, which adds to the other pressures (i.e., of radiation, baryons, DM and vacuum pressure) in the total energy–momentum tensor. As should be expected, the creation pressure is defined in terms of the creation rate and other physical quantities. In the case of adiabatic creation of DM, it is given by [46–50,52]

$$p_c = -\frac{\rho_m}{3H}\Gamma. \tag{9}$$

It is clear from the above equation that if the creation pressure p_c is negative, then it may drive the era of cosmic acceleration of the Universe. We have $p_c < 0$, when $\Gamma > 0$ since $\rho_m > 0$ and $H > 0$, i.e. $\dot{a} > 0$, and $p_c > 0$ when $\Gamma < 0$. Let us now assume the particle creation rate Γ to discuss the RDE model. The most natural choice is taken to be [51]

$$\Gamma = 3\beta H, \tag{10}$$

where β is a constant parameter which lies in the interval $0 \leq \beta < 1$. It has been found that this form of Γ does not favour the epoch in the evolution of the Universe. Lima *et al* [53] have investigated the CDM model with this form of Γ, but time-dependent dimensionless parameter β. However, we restrict ourself with β as a constant but study with the other dissipative mechanism, i.e., viscous fluid to find the epoch of the evolution of the Universe. We consider that these two dissipative mechanisms have independent physical phenomena.

Now, using (10) into (9), we obtain

$$p_c = -\beta\rho_m. \tag{11}$$

Using (10) into (8), we find the solution for the particle number density as

$$n = n_0 \left(\frac{a_0}{a}\right)^{3(1-\beta)}, \tag{12}$$

where n_0 is the constant of integration and considered as the present value of particle number density. In adiabatic particle production, the particles and entropy are produced in the space–time, but the specific entropy (per particle) $\sigma = S/N$, remains constant. This implies that $\dot{S}/S = \dot{N}/N$.

Using (1) and (11), a single evolution equation from (6) and (7) can be obtained as

$$[2 + 3\alpha(\beta + \omega_d)]\dot{H} + 3[1 - \beta + 2\alpha(\beta + \omega_d)] H^2 - 3\zeta H = 0. \tag{13}$$

In the absence of bulk viscosity and matter creation, i.e., taking $\zeta = 0$ and $\beta = 0$, the solution of (13) for RDE model is given by

$$H = H_0 \left(\frac{a_0}{a}\right)^{l}, \tag{14}$$

where a_0 and H_0 are respectively the present-day values of the scale factor and Hubble parameter at time t_0 and $l = 3(1 + 2\alpha\omega_d)/(2 + 3\alpha\omega_d)$. The solution of the scale factor is obtained as

$$a = a_0 \left[1 + lH_0(t - t_0)\right]^{1/l}. \tag{15}$$

The scale factor shows the power-law expansion. The deceleration parameter, which is defined as $q = -a\ddot{a}/\dot{a}^2$, gives $q = (l - 1)$. Thus, the value of q is constant. Therefore, the RDE model itself does not describe the transition redshift.

Further, let us consider the RDE model with matter creation in the absence of bulk viscosity, i.e., $\zeta = 0$. The solution of (13) in the presence of matter creation is given by

$$H = H_0 \left(\frac{a_0}{a}\right)^{k}, \tag{16}$$

where $k = 3[1 - \beta + 2\alpha(\beta + \omega_d)]/[2 + 3\alpha(\beta + \omega_d)]$. The solution of the scale factor is given by

$$a = a_0 \left(1 + kH_0(t - t_0)\right)^{1/k}. \tag{17}$$

Equation (17) again shows that the expansion of the Universe is of power-law form. The constant value of $q = (k - 1)$ shows that the model does not transit from decelerated phase to accelerated phase. The model decelerates for $k > 1$, has marginal inflation at $k = 1$ and accelerates for $k < 1$. Therefore, the form of Γ defined in eq. (10) does not explain the present-day Universe, i.e., the transition phase in the absence of bulk viscosity.

In what follows, we find the solution of eq. (13) with non-zero bulk viscosity and matter creation by assuming suitable form of bulk viscous coefficient, ζ.

3. Solution with viscosity and matter creation

Following [72–74], we consider ζ of the following form:

$$\zeta = \zeta_0 + \zeta_1 H, \tag{18}$$

where ζ_0 and ζ_1 are positive constants. Equation (18) represents the general assumption as it is a combination of two forms, $\zeta = \zeta_0$ and $\zeta \propto H$. This motivation can be traced in fluid mechanics where the transport/viscosity phenomenon is involved with velocity $\dot{a}$ which is related to the expansion rate H. We define the dimensionless bulk viscous coefficients ξ, ξ_0 and ξ_1 as

$$\xi = \frac{\zeta}{H_0}, \quad \xi_0 = \frac{\zeta_0}{H_0} \quad \text{and} \quad \xi_1 = \zeta_1, \tag{19}$$

where H_0 is the current value of the Hubble parameter. Using the above transformation in (18), we obtain the dimensionless form of bulk viscosity as $\xi = \xi_0 + \xi_1 h$, where $h = H/H_0$ is the dimensionless Hubble parameter.

Using the relation

$$\frac{\mathrm{d}}{\mathrm{d}t} = \frac{\dot{a}}{a}\frac{\mathrm{d}}{\mathrm{d}\ln a}$$

in (13), a dimensionless evolution equation for $h = H/H_0$ is given by

$$\psi_2 h' + 3\psi_1 h - 3\xi_0 = 0, \tag{20}$$

where $\psi_1 = [1 - \beta - \xi_1 + 2\alpha(\beta + \omega_d)]$, $\psi_2 = [2 + 3\alpha(\beta + \omega_d)]$ and the prime denotes the differentiation with respect to the conformal time $\ln a$.

Integration of eq. (20) gives

$$h(a) = \frac{\xi_0}{\psi_1} + \left(1 - \frac{\xi_0}{\psi_1}\right)\left(\frac{a}{a_0}\right)^{-3\psi_1/\psi_2}, \qquad \psi_1 \neq 0. \tag{21}$$

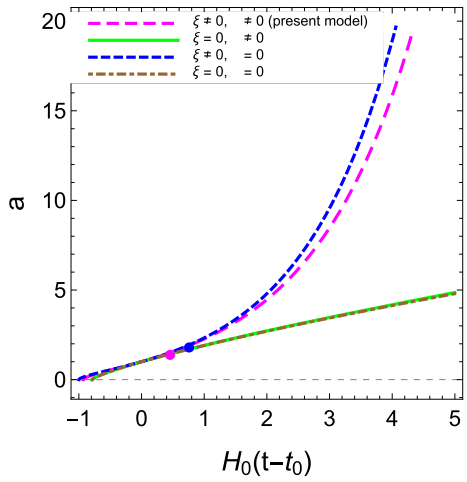


Figure 1. The evolution of the scale factor with respect to $H_0(t - t_0)$ for the best-fit values of free parameters. The dot denotes the transition value. The brown and green trajectories of the scale factor show decelerated expansion whereas the blue and magenta trajectories show the accelerated expansion after the transition point.

Now, using the relation $(a_0/a) = (1 + z)$, we can write the Hubble parameter H in terms of redshift z as

$$H(z) = H_0\left[\frac{\xi_0}{\psi_1} + \left(1 - \frac{\xi_0}{\psi_1}\right)(1 + z)^{3\psi_1/\psi_2}\right]. \tag{22}$$

We derive the t–z relationship, which comes out as

$$t(z) = t_0 + \frac{\psi_2}{3\xi_0 H_0}\ln\left[1 + \frac{\xi_0}{\psi_1}\{(1 + z)^{-3\psi_1/\psi_2} - 1\}\right]. \tag{23}$$

From (22), the scale factor $a(t)$ is obtained as

$$a(t) = a_0\left[1 + \frac{\psi_1}{\xi_0}\left(\mathrm{e}^{\frac{3\xi_0 H_0(t-t_0)}{\psi_2}} - 1\right)\right]^{\psi_2/3\psi_1}, \quad \xi_0 \neq 0, \quad \psi_1 \neq 0, \tag{24}$$

where a_0 is the scale factor at the present epoch and now onwards can be taken as 1. To better understand the evolution of the scale factor, we have plotted the graph of $a(t)$ with respect to $H_0(t - t_0)$ in figure 1 for the best-fitted values of the model parameters. The best-fit values of the model parameters from the observations are given in table 1 (discussed in §4). From figure 1, it is observed that the accelerated expansion starts earlier in the RDE model with bulk viscosity and matter creation (violet trajectory) than the RDE model with bulk viscosity (blue trajectory) during the evolution of the Universe.

The Hubble parameter $H(t)$ in terms of t can be written as

$$H(t) = H_0\mathrm{e}^{\frac{3\xi_0 H_0(t-t_0)}{\psi_2}}\left[1 + \frac{\psi_1}{\xi_0}\left(\mathrm{e}^{\frac{3\xi_0 H_0(t-t_0)}{\psi_2}} - 1\right)\right]^{-1}.$$

Table 1. The best-fit values of the free parameters of RDE model with bulk viscosity and matter creation using SNe + OHD + BAO/CMB samples

Parameters	Best-fit values
ξ_0	5.223
ξ_1	0.485
β	0.777
ω_d	−0.240
α	8.078
M	24.95
$\chi^2_{\min}$	58.79
χ^2_{red}	0.77

(25)

The effective dark energy density, ρ_{eff}, and effective pressure, p_{eff}, are respectively given as

$$\rho_{\text{eff}} = 3H_0^2 \mathrm{e}^{\frac{6\xi_0 H_0(t-t_0)}{\psi_2}} \left[1 + \frac{\psi_1}{\xi_0}\left(\mathrm{e}^{\frac{3\xi_0 H_0(t-t_0)}{\psi_2}} - 1\right)\right]^{-2} \tag{26}$$

and

$$p_{\text{eff}} = -\frac{3H_0^2\xi_0 \mathrm{e}^{\frac{3\xi_0 H_0(t-t_0)}{\psi_2}}\left[\psi_2 \mathrm{e}^{\frac{3\xi_0 H_0(t-t_0)}{\psi_2}} + 2(\xi_0 - \psi_1)\right]}{\psi_2\left[1 + \frac{\psi_1}{\xi_0}(\mathrm{e}^{\frac{3\xi_0 H_0(t-t_0)}{\psi_2}} - 1)\right]^2}. \tag{27}$$

4. Parameters estimation

We use the Hubble parameter obtained in eq. (22) and estimate the best fit of the model parameters $\xi_0, \xi_1, \beta, \omega_d$ and α using the combined data set, consisting of the Type Ia supernova (SNe) data set, observational Hubble data (OHD) and the combined baryon acoustic oscillation (BAO) and cosmic microwave background (CMB) data. In order to figure out the observational constraints, we employ the publicly available EMCEE codes [75] for implementing the Markov chain Monte Carlo (MCMC) method. We choose the current Hubble constant value as $H_0 = 67.8$ km/s/Mpc from Planck 2015 results [76].

It is believed that the SNe observation provides the evidence of cosmic accelerating expansion. For this purpose, we used 31 binned data points of the cJLA data set with the range $0.01 < z < 1.3$ [77] for which χ^2 is defined by

$$\chi^2_{\text{SNe}} = r^t C_b^{-1} r, \tag{28}$$

where $r = \mu_b - M - 5\log_{10} d_L$. Here μ_b is the observational distance modulus for a supernova at redshift z, M is a free normalisation parameter and C_b is the covariance matrix of μ_b for which see table F.2 in [77]. In the above equation, d_L is the luminosity distance which is defined by

$$d_L = \frac{c(1+z)}{H_0}\int_0^z \frac{\mathrm{d}z'}{H(z',\theta)}, \tag{29}$$

where c is the speed of light and θ is the representation of model parameters α, β, ω_d, ξ_0 and ξ_1.

In the recent past, the measurement of Hubble parameter $H(z)$ got much attention from the researchers due to its model-independent nature. In this paper, we compare our model with 43 data points of the Hubble parameter in the red-shift range $0 < z < 2.5$ [78]. The χ^2 function is defined as follows:

$$\chi^2_{\text{OHD}} = \sum_{i=1}^{n} \frac{[H(z_i) - H_{\text{obs}}(z_i,\theta)]^2}{\sigma_i^2}, \tag{30}$$

where $H(z_i)$ and $H_{\text{obs}}(z_i,\theta)$ represent the theoretical and observed values of Hubble parameter, respectively. The standard deviation in the observed value is denoted by σ_i.

We consider BAO and CMB data, i.e., BAO/CMB data from different observational missions [79]. We combine the sample of BAO distance measurements from SDSS(R) [80], isotropic BAO measurement of 6dF Galaxy survey [81], BOSS CMASS [82] and three parallel measurements from WiggleZ survey [83] as the latest observational data for BAO. We combine these results with the Planck 2015 results [84].

In the context of BAO, the angular diameter, $d_A(z,\theta)$ is defined as

$$d_A(z_*,\theta) = c\int_0^{z_*} \frac{\mathrm{d}z'}{H(z',\theta)}, \tag{31}$$

where z_* indicates the photon decoupling redshift and holds the value $z_* = 1090$ as per the Planck 2015 results [84]. $D_v(z,\theta)$ represents the dilation scale which is given by

$$D_v(z,\theta) = \left(\frac{d_A^2(z,\theta)cz}{H(z,\theta)}\right)^{1/3}.$$

The distant redshift d_z is given by

$$d_z = \frac{r_s(z_*)}{D_v(z)}, \tag{32}$$

where $r_s(z_*)$ is defined as the co-moving sound horizon at the time when photons decouple, which is assumed to be the same as it is considered in ref. [79].

The χ^2 function can be explained as [79]

$$\chi^2_{\text{BAO/CMB}} = A^T C^{-1} A, \tag{33}$$

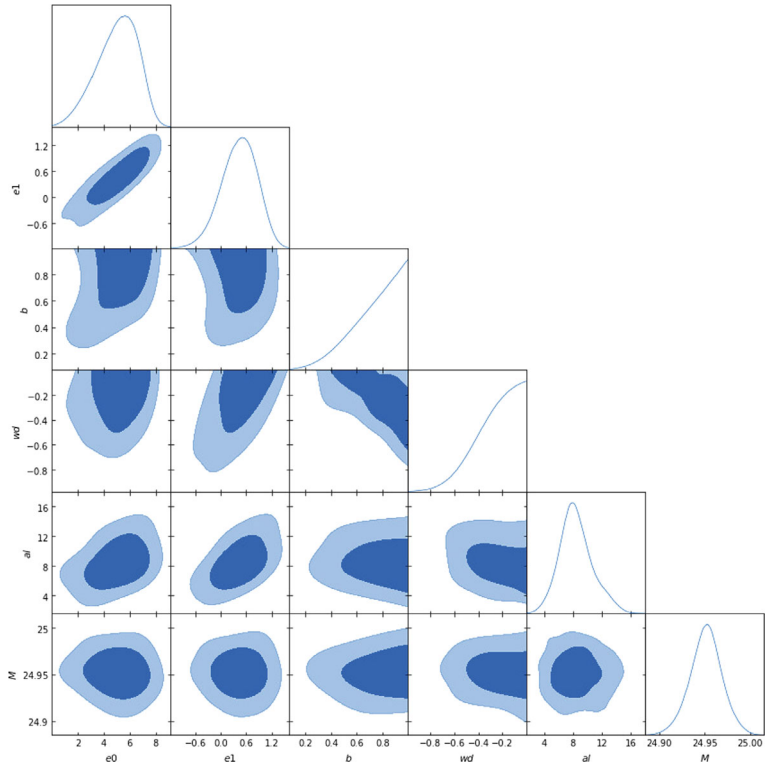


Figure 2. The contour plot for the free parameters using the observational data SNe + OHD + BAO/CMB for RDE model with bulk viscosity and matter creation. The labels $e0$, $e1$, b, wd and al denote ξ_0, ξ_1, β, ω_d and α parameters, respectively.

where A is the matrix

$$A = \begin{bmatrix} \dfrac{d_A(z_*,\theta)}{D_v(0.106,\theta)} - 30.84 \\ \dfrac{d_A(z_*,\theta)}{D_v(0.35,\theta)} - 10.33 \\ \dfrac{d_A(z_*,\theta)}{D_v(0.57,\theta)} - 6.72 \\ \dfrac{d_A(z_*,\theta)}{D_v(0.44,\theta)} - 8.41 \\ \dfrac{d_A(z_*,\theta)}{D_v(0.6,\theta)} - 6.66 \\ \dfrac{d_A(z_*,\theta)}{D_v(0.73,\theta)} - 5.43 \end{bmatrix}$$

C^{-1} is the inverse of the covariance matrix [79]. We have taken the correlation coefficient from ref. [85].

Considering these cosmological data sets together, i.e. (SNe + OHD + BAO/CMB), the total χ^2 function is then given by

$$\chi^2 = \chi^2_{\rm SNe} + \chi^2_{\rm OHD} + \chi^2_{\rm BAO/CMB}. \tag{34}$$

The best-fit values for the model parameters obtained by minimising total χ^2 are given in table 1 and the contour plot is shown in figure 2 with 1σ (68.3%) and 2σ (95.4%) confidence level.

5. Evolution of cosmological parameters

The Hubble parameter in eq. (21) shows a decreasing behaviour with the scale factor. It can have infinitely large value in the early stages and decreases as the Universe expands and finally saturated to a constant value as $a \to \infty$. Figure 1 shows the behaviour of the scale factor with cosmic time. The transition point is found to be $a_{\rm tr} = 0.936$ where the model transits from decelerated phase to accelerated phase.

The deceleration parameter q is a geometric parameter which measures the state of acceleration/deceleration of the Universe. The positive value of q represents the decelerated phase of the Universe whereas the negative value represents the accelerated phase. Using (24), we obtain

$$q(t) = -1 + \frac{3(\psi_1 - \xi_0)}{\psi_2} e^{\frac{-3\xi_0 H_0 (t-t_0)}{\psi_2}}. \tag{35}$$

Thus, the deceleration parameter (DP) in terms of z can be obtained as

$$q(z) = -1 + \frac{3(\psi_1 - \xi_0)}{\psi_2} \left[\frac{\psi_1}{\psi_1 + \xi_0\{(1+z)^{-3\psi_1/\psi_2} - 1\}} \right]. \tag{36}$$

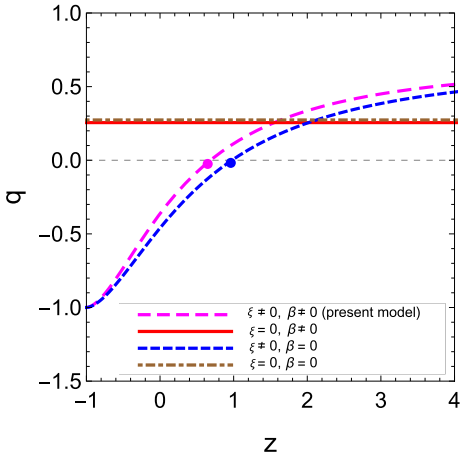


Figure 3. The q–z relation diagram for the best-fitted model parameters. The dot denotes the transition point from where the model transits from decelerated phase to accelerated phase. The horizontal orange and grey trajectories show the constant deceleration values whereas the transition trajectories (blue and violet) are due to the RDE model with bulk viscosity, and the RDE model with bulk viscosity and matter creation, respectively.

The present value of deceleration parameter, q_0, corresponding to $z = 0$ is given by

$$q_0 = -1 + \frac{3(\psi_1 - \xi_0)}{\psi_2}. \tag{37}$$

For best-fit value of parameters, $q_0 = -0.362$, is higher than the corresponding WMAP value $q_0 = -0.60$ [86]. In figure 3 we plot the deceleration parameter $q(z)$ against cosmic redshift z for best fitted values of RDE model. It is observed that q changes its sign from positive to negative showing the transition from decelerated phase to accelerated phase. The transition redshift, $z_{\rm tr}$, is found to be $z_{\rm tr} = 0.68$ for the RDE model, which is within the range of the transition redshift $z_{\rm tr} = 0.45$–0.73 in the concordance ΛCDM model [87].

We can obtain the effective equation of state (EoS) parameter, $\omega_{\rm eff}$, using the standard relation

$$\omega_{\rm eff} = -1 - \frac{1}{3}\frac{2a}{h}\frac{{\rm d}h}{{\rm d}a}$$

as

$$\omega_{\rm eff} = -1 + \frac{2(\psi_1 - \xi_0)}{\psi_2}\left[\frac{\psi_1}{\psi_1 + \xi_0\{(1+z)^{-3\psi_1/\psi_2} - 1\}}\right]. \tag{38}$$

It can be observed that in the late-time as $z \to -1$, $\omega_{\rm eff} \to -1$ which can also be verified from figure 4. The late-time value $\omega_{\rm eff} = -1$ implies that our model converges to ΛCDM model in future. It is also observed that the RDE model does not cross phantom divide line and thus is free from big-rip singularity.

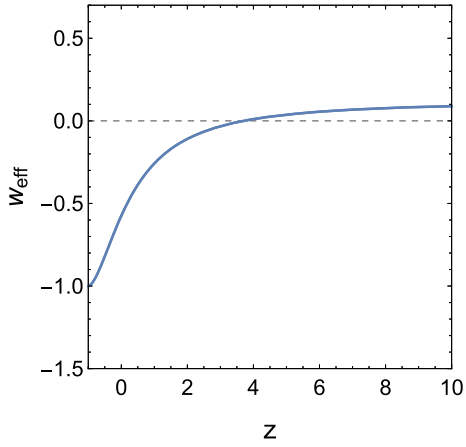


Figure 4. The $\omega_{\rm eff}$–z relation diagram for the best-fitted model parameters.

The present value of $\omega_{\rm eff}$ can be found as

$$\omega_{\rm eff}(z = 0) = -1 + \frac{2(\psi_1 - \xi_0)}{\psi_2}. \tag{39}$$

The present value $\omega_{\rm eff}(z = 0) = -0.575$ can be calculated for the best-fit values of parameters. This value is comparatively higher than that predicted by the joint analysis of WMAP + BAO + H_0 + SNe data which is around -0.93.

To calculate the age of the Universe in terms of redshift, we use the definition given by $t(z) = T(z)/H_0$, where

$$T(z) = \int_z^\infty \frac{{\rm d}z'}{(1+z')(H(z')/H_0)}. \tag{40}$$

For the ΛCDM model, the age parameter is [29]

$$T(z) = \int_z^\infty \frac{{\rm d}z'}{(1+z')[\Omega_{m0}(1+z')^3 + (1-\Omega_{m0})]^{1/2}}. \tag{41}$$

We can calculate and plot the age of the Universe with respect to the redshift z using eq. (22) in (40) for the best-fit values as shown in figure 5. The current age of the Universe is found to be $t_0 = 13.397$ Gyr corresponding to the combined SNe + OHD + BAO/CMB data set, which is comparatively the same as predicted by the ΛCDM model.

The fitting achieved from our statistical analysis for the combined data SNe + OHD + BAO/CMB is compatible with the Hubble observational data for the RDE model as shown in figure 6.

Also, the reduced $\chi^2_{\rm red} = \chi^2_{\rm min}/(N - d)$, where N is the number of data and d is the degree of freedom, is calculated for our model and found to be $\chi^2_{\rm red} = 0.77$ which shows that our RDE model provides a very good fit to the considered observational data.

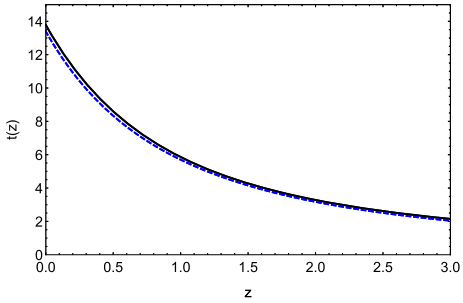


Figure 5. The age of the Universe as a function of redshift where black line represents the age of the ΛCDM model whereas blue dashed line represents the age of the RDE model.

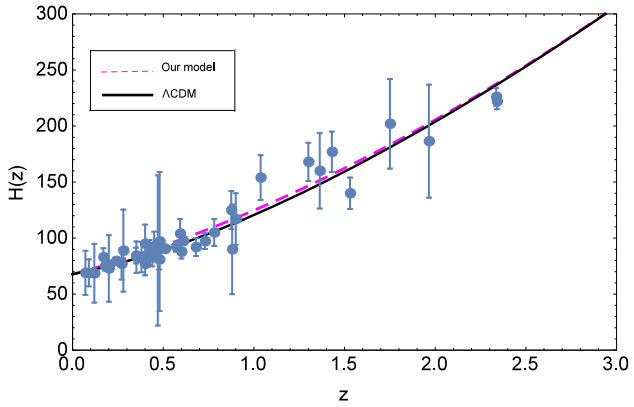


Figure 6. The best-fit curves for the RDE model (dotted magenta) and the ΛCDM model (grey line). The navy blue points with uncertainty bars correspond to the OHD sample.

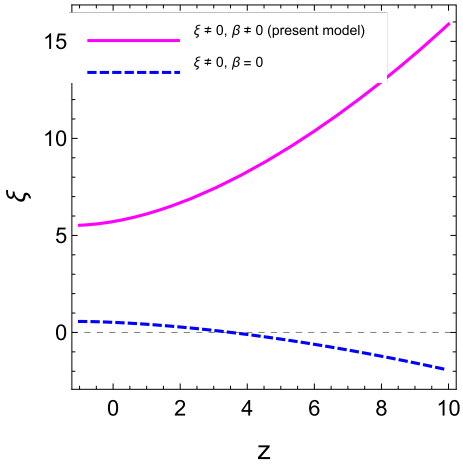


Figure 7. The behaviour of bulk viscous coefficient with redshift for best-fit values of model parameters.

Using (22), we obtain the evolution of bulk viscosity as

$$\xi = \xi_0 + \xi_1\left[\frac{\xi_0}{\psi_1} + \left(1 - \frac{\xi_0}{\psi_1}\right)(1+z)^{3\psi_1/\psi_2}\right]. \quad (42)$$

The trajectory of bulk viscous coefficient for the best-fit values of model parameters (see table 1) is shown in figure 7. It can be observed that the bulk viscosity is negative at higher redshift (early time) and positive in lower redshift (late time) in the RDE model with bulk viscosity (dotted blue curve). This means that the rate of entropy production is negative in the early epoch and positive in the later epoch. Hence, the entropy law violates in the early epoch and obeys in the later epoch. However, the bulk viscosity decreases at low redshift but always positive in the RDE model with bulk viscosity and matter creation (solid magenta curve). Thus, the model does not violate the law of entropy.

6. Geometrical diagnostics

An important tool for investigating DE model characters nowadays is by the introduction of some geometry quantities, for instance the statefinder diagnostic parameters. Sahni *et al* [88] and Alam *et al* [89] introduced a useful geometrical diagnostic pair $\{r, s\}$, called 'statefinder parameter', which is defined as

$$r = \frac{\dddot{a}}{aH^3} \text{ and } s = \frac{r-1}{3(q-1/2)}, \quad (43)$$

where r is the jerk parameter and s is a function of the jerk and the decelerating parameter q. The statefinder parameter pair $\{r, s\}$ of the model concerned is calculated explicitly to demonstrate the behaviour of the RDE model with bulk viscosity and matter creation which is obtained as

$$r = 1 + \frac{9(\xi_0-\psi_1)\left(1-\frac{\psi_1}{\psi_2}\right)}{\psi_2 e^{3\xi_0 H_0(t-t_0)/\psi_2}} + \frac{9(\xi_0-\psi_1)^2}{\psi_2^2 e^{6\xi_0 H_0(t-t_0)/\psi_2}} \quad (44)$$

$$s = \frac{\frac{2(\xi_0-\psi_1)\left(1-\frac{\psi_1}{\psi_2}\right)}{\psi_2 e^{\frac{3\xi_0 H_0(t-t_0)}{\psi_2}}} + \frac{2(\xi_0-\psi_1)^2}{\psi_2^2 e^{\frac{6\xi_0 H_0(t-t_0)}{\psi_2}}}}{\frac{2(\psi_1-\xi_0)}{\psi_2 e^{\frac{3\xi_0 H_0(t-t_0)}{\psi_2}}} - 1}. \quad (45)$$

By using (44) and (45), we plot the r–s plane in figure 8 which shows that the trajectory starts evolving from the region $r > 1, s < 0$, which represents the behaviour of DE with chaplygin gas and evolves through the region $r < 1, s > 0$, which represents the quintessence model of DE and eventually approaches to the point of ΛCDM model represented by the point $\{1, 0\}$. Thus, the RDE model with bulk viscosity and matter creation discriminates with the ΛCDM model in early time. However, the model behaves like the ΛCDM model in the late-time evolution of the Universe.

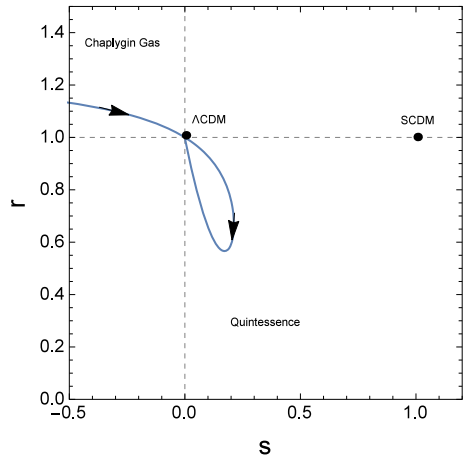


Figure 8. The Universe evolution is diagnostic by the evolution trajectory in the r–s plane for the best-fitted RDE model. The fixed point (0, 1) corresponds to the ΛCDM model. The arrow shows the direction of the evolution of the trajectory.

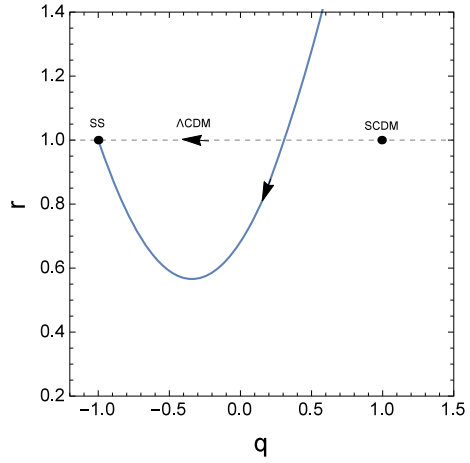


Figure 9. Evolution trajectory in the r–q plane to demonstrate the Universe evolution for the best-fitted RDE model. The horizontal dotted line indicates the ΛCDM model evolution trajectory. The arrow shows the direction of the evolution of the trajectory.

To illustrate more details for the RDE model which is different from the ΛCDM model, we also plot the function trajectory in the r–q plane as shown in figure 9. The dashed line shows the time evolution of the ΛCDM model. The dashed line of the ΛCDM model divides the plane into two parts in which the upper part represents the phantom evolution of the model whereas the lower part represents the non-phantom phase. We can see that the model shows phase transition from deceleration to acceleration as q changes its sign from positive to negative.

To examine the dynamics of the DE models one more diagnostic, called Om, is introduced in [90] with the help of Hubble parameter $H(z)$ and redshift z. It is

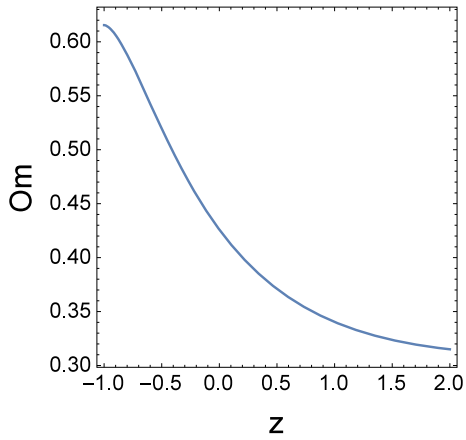


Figure 10. The $Om(z)$–z diagram for the RDE model corresponding to the best-fit parameters.

defined as

$$Om(z) = \frac{(H^2(z)/H_0^2) - 1}{(1+z)^3 - 1}. \tag{46}$$

The Om-diagnostic involves only the first-order derivative of the scale factor. Thus, it is easier to reconstruct from the observational data. The Om-diagnostic is used to check different periods of the Universe. We plot the trajectory in $Om(z)$–z plane to discriminate the behaviour of the DE models. The positive slope of the Om diagnostic means that the model behaves phantom-like while the negative slope indicates the behaviour of quintessence model.

For the RDE model, we obtain

$$Om(z) = \frac{[\xi_0 + (\psi_1 - \xi_0)(1+z)^{3\psi_1/\psi_2}]^2 - \psi_1^2}{\psi_1^2[(1+z)^3 - 1]}. \tag{47}$$

From figure 10, we can observe that the negative slope of the trajectory of $Om(z)$–z plane indicates a behaviour similar to the quintessence model.

To study more about the evolution of the Universe, we expand the scale factor $a(t)$ at $t = t_0$ using Taylor series expansion as

$$a(t) = \sum_{i=0}^{\infty} \frac{a^{(i)}(t_0)}{i!}(t - t_0)^i$$

which provides us some geometrical parameters, known as cosmographic parameters (CP), such as jerk j, snap s and lerk l parameters which can be respectively defined as [91,92]

$$j = \frac{1}{aH^3}\frac{\mathrm{d}^3 a}{\mathrm{d}t^3}, \quad s = \frac{1}{aH^4}\frac{\mathrm{d}^4 a}{\mathrm{d}t^4}, \quad l = \frac{1}{aH^5}\frac{\mathrm{d}^5 a}{\mathrm{d}t^5}. \tag{48}$$

Here r and j are the same; but s defined in (48) is not the same with the one defined in (43). For our model, the variations of parameters j, s and l are discussed by

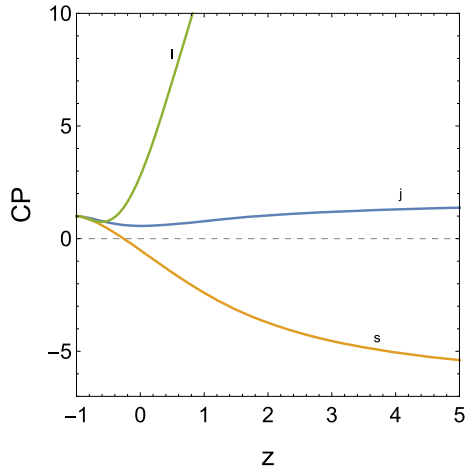


Figure 11. The evolution of the cosmographic parameters jerk j, snap s and lerk l for the best-fit model parameters.

plotting their trajectories against the redshift z as shown in figure 11.

From figure 11, we notice that the jerk parameter remains positive throughout and converges to 1 as z tends to -1 which corresponds to the ΛCDM model. This may indicate that at the present time our model is different from the ΛCDM model but at the late time it converges to the ΛCDM model. The trajectory of s shows the transition from negative to positive, i.e. at the early time s takes the negative value whereas it becomes positive and approaches 1 in the late-time evolution. The trajectory of l remains positive without any transition throughout the evolution and also approaches 1.

7. Energy conditions

In general theory of relativity, the study of singularity theory of space–time was based on energy conditions (ECs). The ECs take the form of various linear combinations of the stress-energy tensor components in such a way that the energy remains positive, or at least non-negative [93]. The famous Raychaudhuri's equation for the expansion nature gives rise to the ECs [94]. Here it is to be noted that the Raychaudhuri's equation is purely geometric and it makes no reference to any gravitational theory under consideration. In FRW Universe, the ECs take the following forms:

$$\text{NEC} \Leftrightarrow \rho_{\text{eff}} + p_{\text{eff}} \geq 0, \tag{49}$$

$$\text{WEC} \Leftrightarrow \rho_{\text{eff}} \geq 0 \ \text{ and } \ \rho_{\text{eff}} + p_{\text{eff}} \geq 0, \tag{50}$$

$$\text{SEC} \Leftrightarrow \rho_{\text{eff}} + 3\, p_{\text{eff}} \geq 0 \ \text{ and } \ \rho_{\text{eff}} + p_{\text{eff}} \geq 0, \tag{51}$$

$$\text{DEC} \Leftrightarrow \rho_{\text{eff}} \geq 0 \ \text{ and } \ \rho_{\text{eff}} \geq |p_{\text{eff}}|, \tag{52}$$

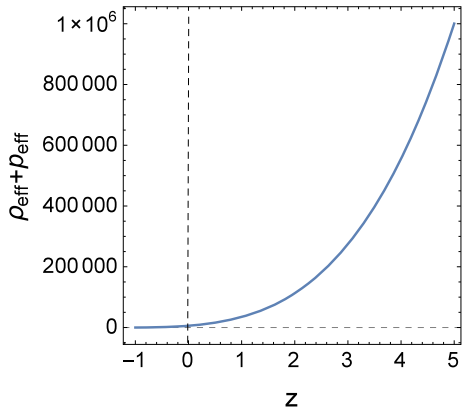


Figure 12. The plot of NEC against z for the best-fit model parameters.

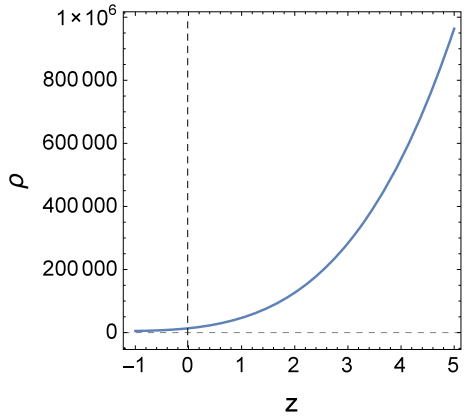


Figure 13. The plot of ρ against z for the best-fit model parameters.

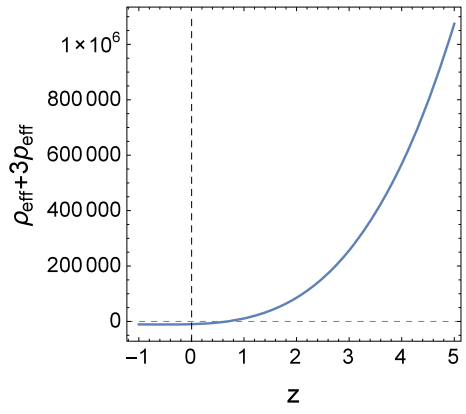


Figure 14. The plot of SEC against z for the best-fit model parameters.

where NEC, WEC, SEC and DEC correspond to the null, weak, strong and dominant energy conditions, respectively. Also in refs [95–97], authors have analysed the ECs in the general theory of relativity.

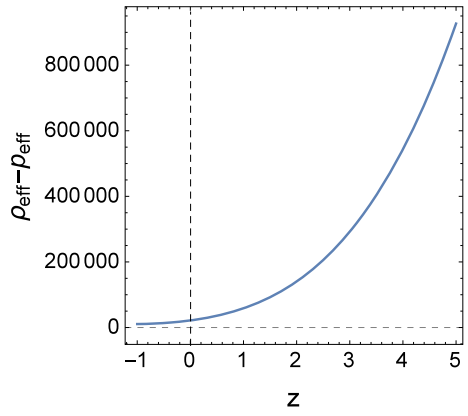


Figure 15. The plot of DEC against z for the best-fit model parameters.

One can observe from figures 12–15 that our model satisfies the NEC, WEC and DEC while it violates the SEC.

8. Conclusion

In the present work, we have discussed the Ricci DE model while considering the effects of bulk viscosity and adiabatic matter creation within the framework of the flat FRW Universe. We have assumed the bulk viscosity coefficient as $\zeta = \zeta_0 + \zeta_1 H$ and the particle creation rate as $\Gamma = 3\beta H$ to obtain exact solutions for the scale factor and various physical quantities. We have considered these two dissipative phenomena as independent irreversible processes. It has been observed that the assumption $\Gamma \propto H$ does not describe the present-day Universe (transition phase) in the absence of bulk viscosity. Also, the RDE model without bulk viscosity and matter creation does not show transition redshift. Therefore, to overcome this problem we have introduced bulk viscosity along with matter creation to observe the present-day Universe and we have succeeded to obtain such a model. We have obtained the Hubble parameter in terms of redshift which is used to obtain the free parameters of the RDE model.

We have performed statistical analysis for our model in §4 using the latest observational data of SNe, OHD and combined data of BAO and CMB. We have employed the publicly available EMCEE codes for implementing the MCMC method. We have obtained the best-fit values for the model parameters (see table 1). Using best-fit values, we have found the transition value of the scale factor ($a_{tr} = 0.936$), the current of the deceleration parameter ($q_0 = -0.362$) and EoS parameter ($\omega_{eff} = -0.575$).

To get further information about the behaviour of the model, we have presented evolution of various cosmological parameters and studied various geometrical diagnostics and cosmographic parameters both analytically and graphically. The deceleration parameter shows a signature flipping behaviour thereby indicating the evolution of the Universe from early deceleration to present late time acceleration (see figure 3). The behaviour of effective EoS parameters (see figure 4) shows that RDE model behaves like de Sitter Universe in late-time of evolution. It has also been observed that the trajectory of ω_{eff} does not cross phantom divide line and thus is free from big-rip singularity. Figure 5 shows the age of the Universe with respect to redshift z for the best-fit values of the model parameters. The age of the Universe is found to be 13.397 Gyr which is very close to that predicted by the ΛCDM model. From figure 6, it has been observed that these best-fit values of the model parameters are in good agreement with the predictions of the Standard Model. The reduced χ^2_{red} obtained for the RDE model provides a very good fit to the considered observational data.

Furthermore, we have performed the statefinder diagnostic analysis to this RDE model with bulk viscosity and matter creation to discriminate from the ΛCDM model. In figure 8, we have plotted the r–s trajectory and observed that our model is finally approaching ΛCDM while in figure 9 there is a sign change of q from positive to negative which shows that our model changes the phase from deceleration to acceleration. In both the figures arrows show the direction of the trajectories in the plane. The Om diagnostic which is obtained from the Hubble parameter shows the negative slope in $Om(z)$–z plane trajectory (figure 10), i.e. its behaviour is similar to the quintessence model. We have also discussed cosmographic parameters like j, s, l to compare our model with the ΛCDM model. From figure 11, we can observe from the trajectory of jerk j that at present our model is different from the ΛCDM model but in the late time it converges to ΛCDM. We can observe from the trajectory of s that it changes its sign from negative to positive during the evolution and finally converges to 1. Similarly, l shows no transition with respect to z and reaches 1 in the late future.

Finally, we have examined the ECs of our model to analyse the physical viability of the model. It has been observed from figures 12–15 that our model satisfies NEC, WEC and DEC but it violates the SEC.

Acknowledgements

One of the author, AK would like to thank University Grants Commission, New Delhi for providing Senior Research Fellowship under UGC-NET Scholarship.

References

[1] A G Riess *et al*, *Astron. J.* **116**, 1009 (1998)
[2] S Perlmutter *et al*, *Nature* **391**, 51 (1998)
[3] S Perlmutter *et al*, *Astrophys. J.* **517**, 565 (1999)
[4] K Bamba, S Capozziello, S Nojiri and S D Odintsov, *Astrophys. Space Sci.* **342**, 155 (2012)
[5] M Li, X D Li, S Wang and Y Wang, *Commun. Theor. Phys.* **56**, 525 (2011)
[6] B Wang, E Abdalla, F Atrio-Barandela and D Pavon, *Rep. Prog. Phys.* **79(5)**, 096901 (2016)
[7] P J E Peebles and B Ratra, *Rev. Mod. Phys.* **75**, 559 (2003)
[8] E J Copeland, M Sami and S Tsujikawa, *IJMPD* **15**, 1753 (2006)
[9] J A Frieman, M S Turner and D Huterer, *Annu. Rev. Astron. Astrophys.* **46**, 385 (2008)
[10] R R Caldwell and M Kamionkowski, *Annu. Rev. Astron. Astrophys.* **59**, 397 (2009)
[11] M Li, X D Li, S Wang and Y Wang, *Front. Phys.* **8**, 828 (2013)
[12] T Padmanabhan, *Phys. Rep.* **380**, 235 (2003)
[13] V Sahni and A Starobinsky, *Int. J. Mod. Phys. D* **9**, 373 (2000)
[14] R Bousso, *Rev. Mod. Phys.* **74**, 825 (2002)
[15] M Li, *Phys. Lett. B* **603**, 1 (2004)
[16] W Fischler and L Susskind, arXiv:hep-th/9806039
[17] Y Gong, *Phys. Rev. D* **70**, 064029 (2004)
[18] B Wang, Y Gong and E Abdalla, *Phys. Lett. B* **624**, 141 (2005)
[19] X Zhang, *Int. J. Mod. Phys. D* **14**, 1597 (2005)
[20] D Pavon and W Zimdahl, arXiv:hep-th/0511053
[21] A G Cohen *et al*, *Phys. Rev. Lett.* **82**, 4971 (1999)
[22] C Gao, F Wu and X Chen, *Phys. Rev. D* **79**, 043511 (2009)
[23] C J Feng, *Phys. Lett. B* **670**, 231 (2008)
[24] M Bouhmadi-López and Y Tavakoli, *Phys. Rev. D* **87**, 023515 (2013)
[25] M Suwa, K Kobayashi and H Oshima, *J. Mod. Phys.* **6**, 327 (2015)
[26] S Ghaffari, A Sheykhi and M H Dehghani, *Phys. Rev. D* **91**, 023007 (2015)
[27] C W Misner, *Astrophys. J.* **151**, 431 (1968)
[28] Y D Xu, Z G Huang and X H Zhai, *Astrophys. Space Sci.* **337**, 493 (2012)
[29] C J Feng and X Z Li, *Phys. Lett. B* **680**, 355 (2009)
[30] S Nojiri and S D Odintsov, *Phys. Rev. D* **72**, 023003 (2005)
[31] S Capozziello *et al*, *Phys. Rev. D* **73**, 043512 (2006)
[32] N Cruz, S Lepe and F Pena, *Phys. Lett. B* **646**, 177 (2007)
[33] J Ren and X H Meng, *Phys. Lett. B* **633**, 1 (2006)
[34] M G Hu and X H Meng, *Phys. Lett. B* **635**, 186 (2006)
[35] P Kumar and C P Singh, *Astrophys. Space Sci.* **357**, 120 (2015)
[36] C P Singh, *Pramana – J. Phys.* **71**, 33 (2008)
[37] C P Singh, S Kumar and A Pradhan, *Class. Quantum Grav.* **24**, 455 (2007)
[38] C P Singh and P Kumar, *Eur. Phys. J. C* **74**, 3070 (2014)
[39] C P Singh and Milan Srivastava, *Eur. Phys. J. C* **78**, 190 (2018)
[40] Milan Srivastava and C P Singh, *Astrophys. Space Sci.* **363**, 117 (2018)
[41] C P Singh and A Kumar, *Eur. Phys. J. Plus* **133**, 312 (2018)
[42] C P Singh and Amit Kumar, *Astrophys. Space Sci.* **364**, 94 (2019)
[43] C P Singh and Simran Kaur, *Astrophys. Space Sci.* **365**, 2 (2019)
[44] I Prigogine *et al*, *Proc. Natl. Acad. Sci. USA* **85**, 7428 (1988)
[45] I Prigogine *et al*, *Gen. Relativ. Gravit.* **21**, 8 (1989)
[46] J A S Lima, M O Calvão and I Waga, in *Frontier physics, Essays in honor of Jayme Tiomno* (World Scientific, Singapore, 1990)
[47] M O Calvão, J A S Lima and I Waga, *Phys. Lett. A* **162**, 223 (1992)
[48] J A S Lima and A S M Germano, *Phys. Lett. A* **170**, 373 (1992)
[49] W Zimdahl and D Pavón, *Mon. Not. R. Astron. Soc.* **266**, 872 (1994)
[50] J A S Lima, A S M Germano and L R W Abramo, *Phys. Rev. D* **53**, 4287 (1996)
[51] J S Alcaniz and J A S Lima, *Astron. Astrophys.* **349**, 729 (1999)
[52] W Zimdahl, D J Schwarz, A B Balakin and D Pavón, *Phys. Rev. D* **64**, 063501 (2001)
[53] J A S Lima, F E Silva and R C Santos, *Class. Quant. Grav.* **25**, 205006 (2008)
[54] T Harko and F S N Lobo, *Phys. Rev. D* **87**, 044018 (2013)
[55] S Basilakos and M Plionis, *Astron. Astrophys.* **507**, 47 (2009)
[56] G Steigman, R C Santos and J A S Lima, *J. Cosmol. Astropart. Phys.* **06**, 033 (2009)
[57] J A S Lima, J F Jesus and F A Oliveira, *J. Cosmol. Astropart. Phys.* **11**, 027 (2010)
[58] S Basilakos and J A S Lima, *Phys. Rev. D* **82**, 023504 (2010)
[59] C P Singh and A Beesham, *Astrophys. Space Sci.* **336**, 469 (2011)
[60] C P Singh, *Astrophys. Space Sci.* **338**, 411 (2012)
[61] S Chakraborty, S Pan and S Saha, *Phys. Lett. B* **738**, 424 (2014)
[62] S Pan and S Chakraborty, *Adv. High Energy Phys.* **15**, 654025 (2015)
[63] J Haro and S Pan, *Class. Quantum Gravit.* **33**, 165007 (2016)
[64] C P Singh and Simran Kaur, *Phys. Rev. D* **100**, 084057 (2019)
[65] C P Singh and Ajay Kumar, *Eur. Phys. J. C* **80**, 106 (2020)
[66] D Pavon, J Gariel and G Le Denmat, *Gen. Relativ. Gravit.* **28**, 573 (1996)
[67] I Brevik and G Stokka, *Astrophys. Space Sci.* **239**, 89 (1996)

[68] C P Singh, *Mod. Phys. Lett. A* **27**, 1250070 (2012)
[69] R Sudarsan and V P Johri, *Gen. Relativ. Gravit.* **26**, 41 (1994)
[70] J Triginer and D Pavón, *Gen. Relativ. Gravit.* **26**, 513 (1994)
[71] C Eckart, *Phys. Rev.* **58**, 919 (1940)
[72] X H Meng, J Ren and M G Hu, *Commun. Theor. Phys.* **47**, 379 (2007)
[73] X H Meng and X Dou, *Commun. Theor. Phys.* **52**, 377 (2009)
[74] A Avelino and U Nucamendi, *J. Cosmol. Astropart. Phys.* **08**, 009 (2010)
[75] D Foreman-Mackey, D W Hogg, D Lang and J Goodman, *PASP* **125**, 306 (2013)
[76] P A R Ade *et al*, *Astron. Astrophys.* **594**, 13 (2016)
[77] M Betoule *et al*, *Astron. Astrophys.* **568**, A22 (2014), arXiv:1401.4064 [astro-ph.CO]
[78] S L Cao *et al*, *The Eur. Phys. J. C* **78**, 170 (2018)
[79] M V dos Santos, R R R Reis and I Waga, *J. Cosmol. Astropart. Phys.* **02**, 066 (2016)
[80] N Padmanabhan *et al*, *Mon. Not. R. Astron. Soc.* **427**, 2132 (2012)
[81] F Beutler *et al*, *Mon. Not. R. Astron. Soc.* **416**, 3017 (2011)
[82] L Anderson *et al*, *Mon. Not. R. Astron. Soc.* **441**, 24 (2014)
[83] C Blake *et al*, *Mon. Not. R. Astron. Soc.* **425**, 405 (2012)
[84] P A R Ade *et al*, Planck 2015 results. XIII. Cosmological parameters, Preprint arXiv:1502.01589 (2015)
[85] G Hinshaw *et al*, *Astrophys. J. Suppl. Ser.* **208**, 19 (2013)
[86] WMAP Collabration: G Hinshaw *et al*, *Astrophys. J. Suppl.* **180**, 225 (2009)
[87] U Alam, V Sahni and A A Starobinsky, *J. Cosmol. Astropart. Phys.* **0406**, 008 (2004)
[88] V Sahni, T D Saini, A A Starobinsky and U Alam, *JETP Lett.* **77**, 201 (2003)
[89] U Alam *et al*, *Mon. Not. R. Astron. Soc.* **344**, 1057 (2003)
[90] V Sahni, A Shaleloo and A A Starobinsky, *Phys. Rev. D* **78**, 103502 (2008)
[91] M Visser, *Class. Quant. Gravit.* **21**, 2603 (2004)
[92] M Visser, *Gen. Relativ. Gravit.* **37**, 1541 (2005)
[93] M Visser and C Barcelo, *COSMO-99*, 98 (2000), arXiv:gr-qc/0001099
[94] S Carroll, *Spacetime and geometry: An introduction to general relativity* (Addison Wesley, Boston, 2004)
[95] J Santos and J S Alcaniz, *Phys. Lett. B* **619**, 11 (2005)
[96] J Santos *et al*, *Phys. Rev. D* **76**, 043519 (2007)
[97] A A Sen and R J Scherrer, *Phys. Lett. B* **659**, 457 (2008)

arXiv:2008.13769v1 [cs.HC] 31 Aug 2020

Vyaktitv: A Multimodal Peer-to-Peer Hindi Conversations based Dataset for Personality Assessment

Shahid Nawaz Khan*, Maitree Leekha†, Jainendra Shukla*, Rajiv Ratn Shah*
*IIIT-Delhi, New Delhi, India
†Delhi Technological University, New Delhi, India
Email: shahid17102@iiitd.ac.in, maitreeleekha_bt2k16@dtu.ac.in,
jainendra@iiitd.ac.in, rajivratn@iiitd.ac.in

***Abstract*—Automatically detecting personality traits can aid several applications, such as mental health recognition and human resource management. Most datasets introduced for personality detection so far have analyzed these traits for each individual in isolation. However, personality is intimately linked to our social behavior. Furthermore, surprisingly little research has focused on personality analysis using low resource languages. To this end, we present a novel peer-to-peer Hindi conversation dataset, *Vyaktitv*[1]. It consists of high-quality audio and video recordings of the participants, with Hinglish[2] textual transcriptions for each conversation. The dataset also contains a rich set of socio-demographic features, like income, cultural orientation, amongst several others, for all the participants. We release the dataset for public use, as well as perform preliminary statistical analysis along the different dimensions. Finally, we also discuss various other applications and tasks for which the dataset can be employed.**

***Keywords*-Multimedia; Dataset; Human Computer Interaction;**

I. Introduction

According to psychologists, the study of the personality involves addressing some of the most exciting queries on human psychology, including analysis of how different aspects of people's lives are related, and how they interact with the environment as social beings [1]. All of it is an attempt towards better understanding and embracing the differences and uniqueness amongst individuals.

Several formal definitions of personality exist in the literature [2]. However, loosely speaking, an individual's personality primarily constitutes their social and emotional behavior, cognition, and decision-making abilities [1]. Analyzing these personal characteristics can help improve our understanding and management of people in a way that would not have been possible otherwise. Applications of personality assessment can be dated back to 1920s, when it was used to supplement the process of personnel recruitment, particularly for armed forces, as well as to identify post-traumatic stress disorder (PTSD) symptoms [3]. Even today, among several others, the National Aeronautics and Space Administration (NASA) uses different kinds of group activities to assess traits that are fundamental for survival in critical situations [4]. Therefore, automatic personality assessment can pave the way for substantial improvements in the aforementioned domains, among several others.

Several datasets have been introduced recently, each adopting a unique method for assessing personality, including the analysis of social media data [5], anonymous essays [6], short video-based self-introductions [7], call-center simulations [8], performance in group discussions [9], etc. While the studies using these datasets present exciting insights into their subjects' personalities, they have the following limitations.

- Firstly, a large proportion of them [5–8, 10–12] miss an essential ingredient for personality assessment: a social environment. Our idiosyncrasies largely stem from the way we routinely interact with those around us [13].
- Another critical aspect is that a large body of work [5, 9, 14] has been devoted to personality assessment where the primary language used was English, with very little literature about the use of other languages. In other words, the advances made so far have not been tested for non-English speakers.

To this end, we propose a novel peer-to-peer interaction-based multimodal dataset for assessment of personality traits. Our method differs from the previously used techniques of group discussions [9], and simulation activities [8] by analyzing individual behavior in an unrestrained setting, where the participants are at liberty to guide their conversations, react openly, and without any formal criteria of judgment that may exist in the former. Furthermore, we assess conversations where the participants interact in Hindi, which is used by over 637 million [3] people around the world. Finally, previous research [15] has shown that factors like age, ethnicity, and educational qualification can significantly impact an individual's personality. We extend these findings by conducting an ethnographic study analyzing several socio-demographic indicators and how they impact an individual's personality. Our contributions are summarized as follows.

- First multimodal peer-to-peer Hindi conversation-based

[1]In Hindi, *Vyaktitv* means *personality*.
[2]Hindi words are written using English alphabets.
[3]Hindi usage statistics from https://www.statista.com/statistics/266808/the-most-spoken-languages-worldwide/

personality assessment dataset. It includes data in three different media forms: (1) video feed of the participants, (2) their audio recordings, and (3) Hinglish transcriptions of the dialogue between the participants.
- Rich socio-demographic indicators for the participants involved in the study. Many of these, like social media usage, public speaking skills, cultural inclination, have not been considered in the past. Our statistical analysis reveals that these factors significantly impact a person's personality.

The rest of this paper has been organized as follows. The next section reviews some of the widely used theoretical models of personality and the techniques used for personality trait detection. Section III describes the data collection process, while Section IV, continues with the statistical analysis of socio-demographic and lexical features, in terms of their relation with different personality traits. Finally, with Section V, we conclude our work and discuss some of the other applications of the rich, multimodal dataset proposed in this work.

II. Literature Review

A. Personality Measures

There are many theories in modern psychology modeling the personality traits of humans based on several interesting measures. We will be discussing some of them briefly. The interested reader may refer to psychological surveys [16] and books [1] for further details on the subject.

One of the oldest ways of modeling personality is the Trait Theory [17], which aims at analyzing different human behaviors by treating them as 'traits, like *Extraversion*, *Emotional Stability*, etc. It usually follows a lexical approach, wherein the subjects use short phrases to discuss their behavior, based on which a score is given for different traits. More recently, Cattell *et al.* [18] proposed the 16PF model, where they used 16 personality factors, including *Dominance*, *Perfectionism*, *Privateness*, and several others, to describe human behavior. Eysenck [19] produced a more concise three-dimensional model for personality assessment, using *Psychoticism*, *Extraversion*, and *Neuroticism* (PEN) as factors. Furthermore, questionnaires are often used to analyze personality traits based on the subjects responses to a set of specific questions. The Myers Briggs Type Indicator (MBTI) [20] is one such test widely used around the world. It is a self-introspective report which reflects on an individuals decision-making process. Specifically, the test classifies the subjects into two categories, along four different dimensions, namely, *Introversion*/*Extraversion*, *Sensing*/*Intuition*, *Thinking*/*Feeling*, *Judging*/*Perception*. Finally, one of the most popular and widely accepted models in the literature, and the one we use as a foundation in our work, is the Five-Factor model, also known as the Big Five model [21]. As in the case of the Trait Theory, which analyses personality differences based on individual behaviors, the Big Five model is based on five personality factors of subjects, summarized in Table I.

Personality Trait	High Score	Low Score
Extraversion (EXT)	Extrovert	Introvert
Neuroticism (NEU)	Anxious	Confident
Agreeableness (AGR)	Trustworthy	Selfish
Conscientiousness (CON)	Disciplined	Confused
Openness (OPN)	Innovative	Unimaginative

Table I: Big Five personality traits, and the interpretations with high and low scores for these traits.

B. Datasets for Personality Assessment

There have been several studies analyzing human personality traits, and they differ substantially in the way they model the personality of individual subjects. Ranging from analyzing telephonic conversations of the participants [8] to assessing these traits based on their resume [22], the experimental designs of these studies vary significantly. Therefore, this section focuses on datasets and experimental methods explored by the researchers to assess personality, and how our *Vyaktitv* dataset differs from them.

Text based analysis. In his study, Wang [5] used social media data for predicting personality by analyzing the relationship between the language used by the users on these platforms and their personality traits. He curated a dataset of around 90,000 users and used their recent posts on Twitter to predict the MBTI [20] personality traits. Many researchers [6, 14] have also analyzed personality traits based on essays written by the subjects. In this regard, a particularly popular benchmark dataset is the one by Pennebaker and King [23], used to predict the Big Five traits based on user-written essays. Automatic prediction of personality impressions based on the verbal content of vlogs has also attracted the research community's active interest [24].

Audio and Vision based analysis. Ivanov *et al.* [8] curated the PersIA dataset, containing a series of telephonic conversations between participants simulating a call center. Furthermore, they also found the participants' verbal backchanneling responses to be useful predictors of their personality. Using vlogs for audio and video-based automatic assessment of personality has also been actively pursued. In this regard, the ChaLearn First Impressions dataset [25] has been widely experimented with.

Early attempts focused on the use of facial expressions to predict personality. The Cohn-Kanade Dataset [26] is a widely used repository for facial expressions, and has been used jointly with other personality-based datasets [11] to predict the traits. In another study, Berkovsky *et al.* [10] followed a different approach by analyzing personality traits based on the subjects' physiological responses to external

stimuli. Specifically, they employed affective videos and images to evoke emotional responses in the participants and used eye-tracking data captured using Eye-Tracking Glasses (ETG) to predict the personality traits. Furthermore, there have also been attempts at predicting the personality traits in crowds. In particular, Favaretto *et al.* [27] developed a video analysis system that could detect the personality, emotion, and cultural aspects of pedestrians using video sequences.

Use of Group Discussions. Several researchers have used group discussions to assess leadership qualities and personality traits. These include the MS-2 [4], the MATRICS [28], and the ELEA-AV [29] corpuses. However, such setups restrain the free flow of ideas, where the discussion is likely to be guided by the leader, and the introverts may get less chance to speak. Moreover, in discussions like these, people want their views to be accepted by the group, and emphasis is on winning [30] (eg: being elected as the leader).

A very few studies, like the one conducted by Kalimeri *et al.* [31], have attempted to model the personality of individuals based on how they interact with their environment in their daily lives. In their experiment, Kalimeri *et al.* used sensors to track the participants' activity over a period of six weeks. In the present study, we follow a different approach to mimic the social setting by making the subjects participate in peer-to-peer conversations and discuss topics from their daily lives. Peer-to-peer conversations have been utilized in a lot of different domains [32, 33]; however, to the best of our knowledge, personality assessment research has not focused on such interactions. Furthermore, in most of the datasets included as a part of the literature, the primary language used was English. Our dataset, *Vyaktitv*, stands as the first Hindi peer-to-peer conversation dataset and sets us apart from the previous studies. Finally, unlike most of these studies, we also collect the socio-demographic data of the participants involved in the study. Factors like gender and age have been known to impact personality traits [15]. We gather several other features (like if subjects have a sibling or not, or if they live with their parents, etc.) and statistically analyze their influence on the Big Five traits.

III. Vyaktitv Dataset

The *Vyaktitv* dataset[4] provides rich multimodal information from peer-to-peer interactions between individuals. The next few subsections discuss the data collection process, its main features, and the annotations done.

A. Data Collection Methodology

Participants. A total of 38 people (24 Male, 14 Female) participated in the experiment. Their average age was 21.47 years, with a standard deviation of 2.25 years. Most of them were undergraduate or PhD students at an Indian university, while some others were employed as researchers and engineers. All the participants were native Hindi speakers, and had formally learnt Hindi through courses at their high school.

[4]Interested researchers can reach out to any of the authors to gain access to the *Vyaktitv* dataset.

[P1]	In context of the Indian cricket, is Virat Kohli a better test-match captain than Mahendra Singh Dhoni?
[P2]	What are your favourite places to visit in India?
[P3]	Do you think pets are an important part of a family?
[P4]	What are your plans after graduation?
[P1]	In your opinion, should cell phones be allowed during class?
[P5]	How do you think global warming is impacting our lives?
[P6]	Should children be detained in high schools?
[P7]	How do you feel about euthanasia (mercy killing)?
[P8]	Do you think watching television can help in developing the minds of young children?
[P9]	In your view, is cloning animals ethical?
[P10]	Comment on the status of education in rural areas, and the high dropout rate among children of under-privileged families.
[P11]	In your opinion, what are the most promising alternative sources of energy, especially for the Indian economy?

Table II: Conversation prompts: Sample topics given to the participants to start a conversation. These prompts were introduced to the participants in Hindi as well.



Figure 1: Experimental Setup: Two separate high quality cameras, and a microphone were used to record the video and audio of the two participants.

Procedure. The participants were called randomly in pairs to interact with each other. Note, that a participant could be friends with some, and a total stranger to others. A random pairing process discarded the possibility of only friendly pairs being recruited for conversations. They were provided with some initial prompts for starting the conversation, as shown in Table II, although they were also encouraged to choose a topic of their liking. Furthermore, they were encouraged to speak in Hindi at all times, but the mode of communication was a mix of Hindi and some bit of English. A total of 25 conversations were recorded, where the average length of each was 16 minutes and 6 seconds, totaling to about 702 minutes of multimodal content.

[Q1] What best describes your gender? (Female, Male)
[Q2] How old are you?
[Q3] Please mention the amount of work experience you have had (in years). Enter 1 if the duration is ¡ 1 year (say, as an intern).
[Q4] What is your current state of employment? (Junior Undergraduate Student, Senior Undergraduate Student, M. Tech. Student, Junior PhD Student, Senior PhD Student, Employed, Unemployed)
[Q5] One a scale of 1 to 3, how would you rate your presence on social media, including platforms like Twitter, Facebook, and Instagram? By "presence", we mean how often do you post and reflect your thoughts on these platforms.
[Q6] Do you have a sibling? (Yes, No)
[Q7] Is your primary area of residence in the city (urban) or somewhere in the outskirts (rural)? (Urban, Rural)
[Q8] Are you currently staying with your parents? (Yes, No)
[Q9] What is your mother's highest qualification? (< Class 10th, Class 10th, Class 12th, Bachelor's/Diploma, Post Graduate/Professional Training)
[Q10] What is your father's highest qualification? (same options as for the previous question)
[Q11] When at home, do you use Hindi as the main language to converse? (Yes, No)
[Q12] Do you think you are a religious/culturally inclined person? (Yes, No)
[Q13] Please select the band that closes matches your family's annual income in INR (₹). (< ₹0.5 million, ₹0.5 million - ₹1 million, > ₹1 million). In USD: (< USD 6600, USD 6600 - USD 13200, > USD 13200)
[Q14] On a scale of 1 to 5, how would rate your public speaking skills, say when formally addressing a group of people, or even with friends?
[Q15] On a scale of 1 to 5, how would rate your writing proficiency?

Table III: Questionnaire for gathering socio-demographic features from the participants involved in the experiment.

B. Multimodal Content

Vyaktitv includes data of the following main categories, collected for both the participants involved in a conversation:

Video and Audio Recordings. We used two time-synchronized 48 MegaPixel cameras to capture the front view of both the participants (Figure 1). A Fifine K669B Condenser Recording USB Microphone (with a sound-to-noise ratio of 78 dB) was used to record the conversation. In addition, the dual mics of a Samsung Galaxy S7 Edge were used in interview mode to record each participant separately. Although the dual mics had active noise cancellation, additional processing was done using the open-source tool Audacity[5] for further noise reduction.

Hinglish Transcriptions. We also provide manual transcriptions of the conversations in Hinglish. The transcribers were recruited from a university and were all native Hindi speakers. They had also received formal education for both Hindi and English at a high-school level.

Transcription Guidelines. The audio files recorded using the dual mics were given to the transcribers with the following instructions:

- Each file had speech signals from two speakers, one of which was more prominent than the other. The annotators were instructed to type the dialogues spoken only by the prominent speaker.
- The language to be used was Hinglish, *i.e.*, Hindi words written using English alphabets.
- Previous research has established the utility of verbal backchannels for analyzing personality [8]. Therefore, we manually annotated our Hinglish transcriptions for several such feedbacks. In particular, the transcribers were instructed to determine and annotate the lexical cues in shown in Table IV. The annotators were introduced to these words and told to annotate any occurrence of such terms by appending "xxx". This made an automatic analysis of these words relatively easy. The transcribers were provided with a sample list of words from each category for reference.

Socio-demographic Information. We collect socio-demographic indicators for all the participants involved in our study[6]. In order to do that, we designed a short questionnaire, shown in Table III. Use of some of these indicators, like age and gender, was inspired by prior work [15], while several others, like if the subjects had a sibling and whether they lived with their family, have not been used previously. Our analysis reveals that these factors significantly impact the subjects' personality traits (see Section IV-A). A comprehensive view for the distributions of these socio-demographic variables is shown in Figure 2.

Personality Traits. Finally, we employ the widely used questionnaire designed by Goldberg [34] to gather the Five-Factor personality traits of the participants. Figure 3 shows the distribution of these scores for the subjects involved in the study.

IV. Data Analysis

A. Analysis of Socio-Demographic Features

Socio-demographic factors like age, level of education, gender, and ethnicity have been known to impact an individual's personality [15]. In this work, we statistically analyze several other indicators (Section III) that have not been considered in the past studies, and can potentially impact personality traits. Specifically, for each pair of a socio-demographic factor and one of the five personality traits, we use a two-sample Kolmogorov-Smirnov (K-S) test under the null hypothesis that the distribution of the scores of that trait conditioned on the indicator's

[5]https://www.audacityteam.org/

[6]Personally identifiable information (PII) has been anonymized to maintain the privacy of the participants.

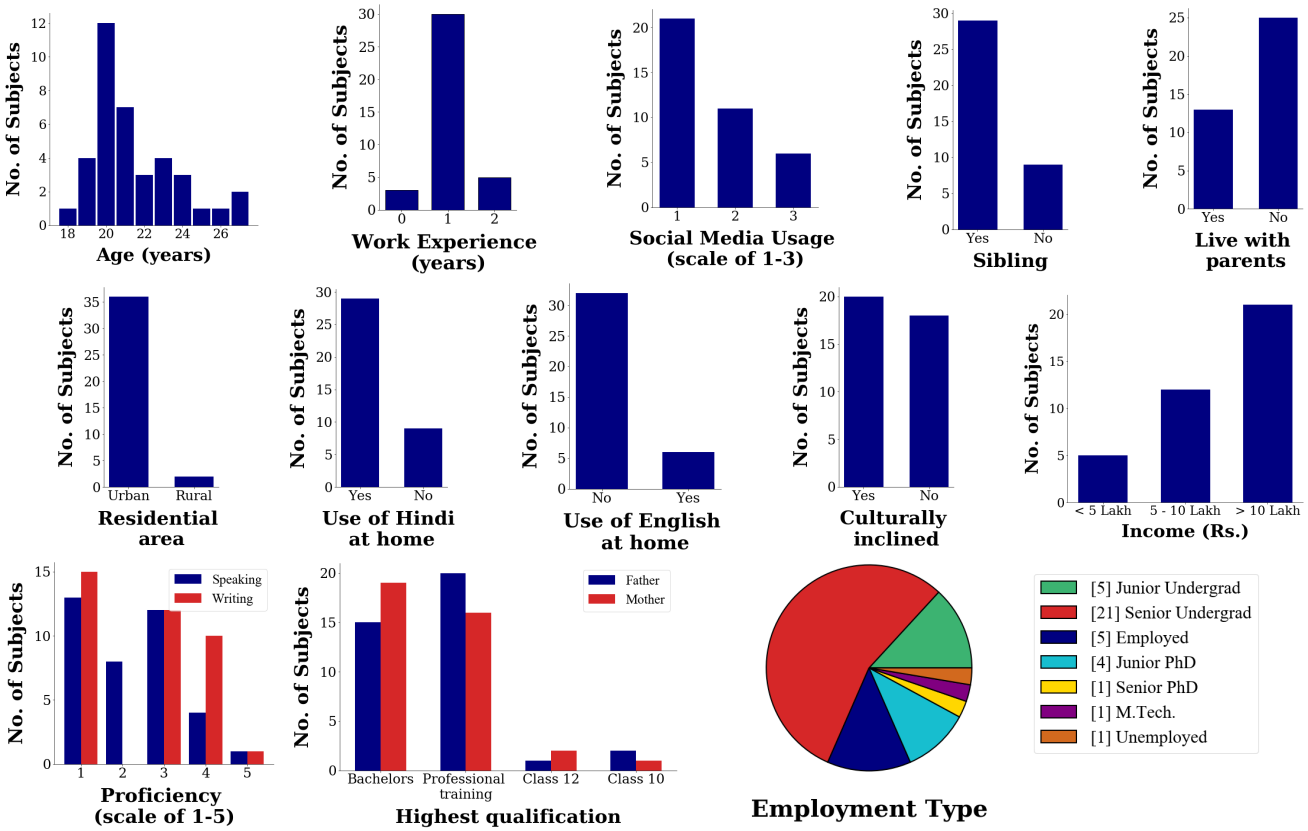


Figure 2: Socio-Demographic Indicators: Distribution of the number of subjects belonging to different categories across these factors.

Lexical Feature	Description	Examples
Filled Pauses	Sounds made by the speaker to fill in the gaps while speaking.	*toh* (so), *aise* (in this way), *hmm*
Discourse Markers	Used to organize the dialogue into segments.	*isliye* (because), *to* (so), *magar* (but), matlab (it means)
Suffix Words	Words that the speaker elongates towards their end.	*areeyy haaan* (ohh yes!), *nahiii* (no), *kyaaa yaar* (what is this)
Repetition of words	Speakers often do this during impromptu conversations.	*jaise... jaise ki* (for example), *magar... magar ye to* (but this is)

Table IV: Lexical Annotations: Examples from the Hinglish transcriptions, along with their closest meaning in English.

values is comparable. Essentially, the test analyzes if there exists a significant difference in the Empirical Cumulative Distribution Functions (ECDF) of a random variable across the two samples under consideration. For instance, an individual may or may not speak in English when at home. Therefore, a two-sample K-S test evaluates the null hypothesis that the distribution of the personality scores for, say *Extraversion*, when the person uses English at home, does not vary significantly from the distribution of the scores when he does not use English at home. For this particular example, however, we shall see shortly that the test rejects the null hypothesis. Since the two-sample K-S test compares only two distributions, we had to restrict the number of unique values of all socio-demographic factors to two, and handle the indicators that had more than two values accordingly. For features like `age`, `work experience`, `social media usage`, `public speaking skills`, and `writing proficiency`, we used the median values of their respective distributions to convert the values to binary, *i.e.*, a value greater than the median would be labelled as `High`, and `Low` otherwise. For other features, like `mother's highest qualification`, `father's highest qualification`, `employment status`, and `income`, we used dummy variables. For instance, `income` $<$ ₹0.5 `million` (˜`USD 6600`) became a separate variable, with a value as 1 if the income was in the range specified, and 0 otherwise. Separate tests were run for all five personality traits.

Table V presents the results of the K-S test conducted for all the Big Five personality traits separately. The features that were found as being of significant importance have been

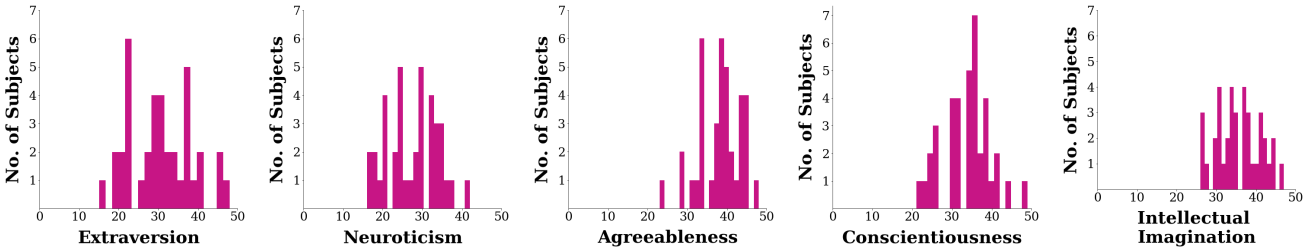


Figure 3: Distribution of the raw Big Five personality trait scores.

Trait	Feature	KS-Statistic
EXT	Father has Professional-training	0.35
	Public Speaking Proficiency	0.39
	Writing Proficiency	0.42
NEU	Cultural	0.36
	Living with parents	0.32
	Social media usage	0.47
AGR	Income $<$ ₹0.5 million	0.48
	Junior Undergrad Student	0.45
	Age	0.29
CON	Social Media usage	0.49
	Sibling	0.28
	Mother's has Professional-training	0.35
OPN	Social Media usage	0.41
	Writing Proficiency	0.41
	Sibling	0.46

Table V: K-S Test: The distribution of personality trait scores, conditioned on the respective socio-demographic features, varied significantly at a 5% significance levels.

shown along with their corresponding value for the KS-Statistic ($\mathcal{K}$). In the interest of brevity, we have tabulated only 3 significantly important indicators per trait. Furthermore, we also plotted the probability distribution functions of the indicators that we found particularly interesting. These plots are shown in Figure 4. It can be seen that the subjects, who had high proficiency in public speaking, were most likely extroverts, as indicated by the distribution of the **EXT** scores. Furthermore, those who were more culturally inclined tended to be more stable emotionally, indicated by lower scores for **NEU**. Participants with annual income less than ₹0.5 million (˜USD 6600) tended to be more straightforward, as depicted by the **AGR** scores. Surprisingly, the **CON** score distribution revealed that subjects who did not have siblings were more disciplined than those who had siblings. Finally, based on the **OPN** distribution, those who had low social media usage were found to be more innovative.

B. Lexical Analysis

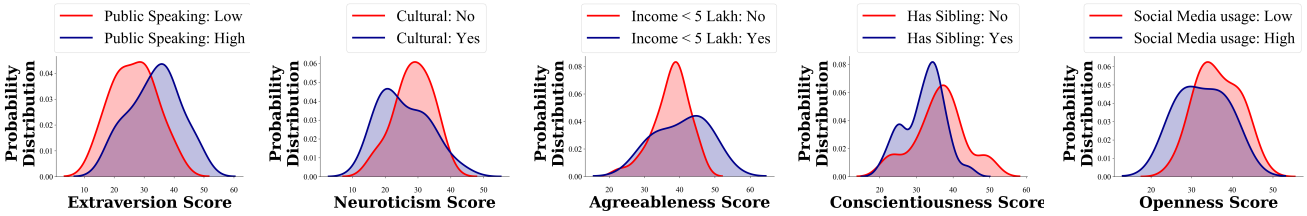
In this section, we present the observations from our analysis of different kinds of lexemes (Table IV) used by the participants during their dialogues. Specifically, we perform a correlation analysis between the Big-Five personality trait scores and the frequency of filled pauses, discourse markers, suffix words, repeating words, and total words used by the participants. The plot in Figure 5 shows the Pearson correlation coefficients between the traits and the frequency of these words, depicting that the overall magnitudes for the correlation coefficients were not significantly high. Following are some other interesting observations we made:

- There is a relatively high correlation between the use of discourse markers and the traits **EXT** and **OPN**. This also aligns with findings of previous studies [35] that established a positive relation between speaker backchannels and high extraversion scores.
- Furthermore, frequency of repetition is positively correlated with **EXT**, and negatively with **NEU**. Low scores for **NEU** indicate high-levels of confidence in the subjects. This is an interesting finding that can be qualitatively analyzed further to draw parallels between individuals' psychological behaviors.
- Subjects who had low scores for **AGR**, *i.e.*, were less cooperative, used relatively more number of suffix words. In addition, these subjects also appeared to make more use of filled pauses, indicating verbal feedback.
- Finally, the total number of words spoken by the subjects in their respective dialogues had a positive relationship with their level of confidence, shown by the negative coefficient for **NEU**. The opposite is true for **AGR**. Participants who used more words tended to have high scores.

V. Recommendations for Future Work

Given the diversity of the *Vyaktitv* dataset, this section outlines some directions for potential future research.

- **Personality Trait Prediction.** A large body of research has been focussed on personality assessment [6, 12, 36]. The dataset includes multimodal features, including the video and audio signals, as well as the transcriptions, which can be used to model personality in a multimodal setting [4, 37]. Since most of the research done so far has been on English datasets, *Vyaktitv* acts as a rich, culturally different dataset in a low resource language that can be used to advance personality assessment



(i) (ii) $(iiii)$ (iv) (v)

Figure 4: Probability Distribution Plots depicting the relation between different socio-demographic indicators and the Big Five personality traits: (i) EXT, (ii) NEU, (iii) AGR, (iv) CON, and (v) OPN.

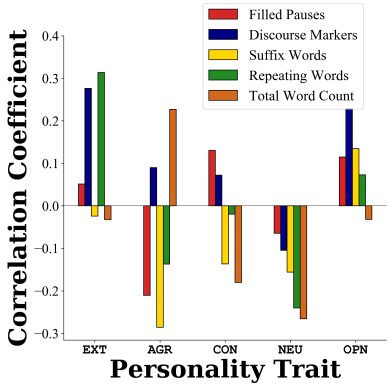


Figure 5: Correlation analysis between the Big-Five personality traits and frequency of filled pauses, discourse markers, suffix words, repeating words, and total words.

research, and tackle challenges peculiar to such low resource languages. For instance, lack of state-of-the-art word embeddings and impact of features that change with population, i.e., progressing towards cross-cultural personality assessment.

- **Modelling and Design of Conversational Agents.** Research on Embodied Conversational Agents (ECA) has long been a primary center of focus for both the AI and HCI communities. Peer-to-Peer conversations have been extensively used for modeling a wide range of tasks that relate to the design of ECAs. These include, but are not limited to: predicting listener backchannels [32, 38], disengagement prediction [32], and speaking turn prediction [39]. With datasets such as *Vyaktitv*, the research in the field of ECA design can potentially progress towards culture-sensitive agents, which can sense and accordingly modify their behavior based on the subjects around.

VI. Conclusion

In this work, we proposed a novel dataset, *Vyaktitv*, for analyzing the personality traits of individuals through peer-to-peer social interactions. It is a multimodal dataset, where the subjects participate in pairwise conversations, and the primary language used is Hindi. We also collected a rich set of socio-demographic indicators from all the participants to study their impact on personality. In particular, we observed that factors like `income`, `public speaking skills`, `whether they live with their parents`, `social media usage`, and several others, were significant in distinguishing the distribution of the personality traits. Observations from the lexical analysis of the Hinglish transcriptions also indicated some interesting findings as to how the use of different types of filler words correlated with personality traits. Towards the end, we highlighted the diversity and richness of the multimodal Hindi dataset proposed in this work by discussing potential research directions that can be pursued. It is not limited to personality assessment, but can also be used for several other challenging tasks that aim at designing human-computer interfaces, like Backchannel Opportunity Prediction (BOP) [38], Listener Disengagement Prediction [32], and modeling conversational agents [40].

Acknowledgements

We would like to acknowledge all the subjects who took part in the experiment, as well as the annotators who assisted in the transcribing the conversations. Furthermore, Jainendra Shukla and Rajiv Ratn Shah are partly supported by the Infosys Center for AI and the Center of Design and New Media at IIIT-Delhi.

References

[1] D. Cervone and L. A. Pervin, *Personality: Theory and Research*.

[2] P. J. Corr and G. Matthews, *The Cambridge handbook of personality psychology*. Cambridge University Press New York, 2009.

[3] Y. Mehta, N. Majumder, A. Gelbukh, and E. Cambria, "Recent trends in deep learning based personality detection," *Artificial Intelligence Review*, pp. 1–27, 2019.

[4] N. Mana, B. Lepri, P. Chippendale, A. Cappelletti, F. Pianesi, P. Svaizer, and M. Zancanaro, "Multimodal corpus of multi-party meetings for automatic social behavior analysis and personality traits detection," *The Medical Roundtable*, pp. 9–14, 2007.

[5] Y. Wang, "Understanding personality through social media," 2015.

[6] X. Sun, B. Liu, J. Cao, J. Luo, and X. Shen, "Who am i? personality detection based on deep learning for texts," in *2018 IEEE International Conference on Communications (ICC)*. IEEE, 2018, pp. 1–6.

[7] J. Yu, K. Markov, and A. Karpov, "Speaking style based apparent personality recognition," in *International Conference on Speech and Computer*. Springer, 2019, pp. 540–548.

[8] A. V. Ivanov, G. Riccardi, A. J. Sporka, and J. Franc, "Recognition of personality traits from human spoken conversations," in *Twelfth Annual Conference of the International Speech Communication Association*, 2011.

[9] F. Pianesi, N. Mana, A. Cappelletti, B. Lepri, and M. Zancanaro, "Multimodal recognition of personality traits in social interactions," in *Proceedings of the 10th International Conference on Multimodal Interfaces*, ser. ICMI 08. New York, NY, USA: Association for Computing Machinery, 2008, p. 5360. [Online]. Available: https://doi.org/10.1145/1452392.1452404

[10] S. Berkovsky, R. Taib, I. Koprinska, E. Wang, Y. Zeng, J. Li, and S. Kleitman, "Detecting personality traits using eye-tracking data," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–12.

[11] M. Gavrilescu, "Study on determining the big-five personality traits of an individual based on facial expressions," in *2015 E-Health and Bioengineering Conference (EHB)*. IEEE, 2015, pp. 1–6.

[12] N. Ahmad and J. Siddique, "Personality assessment using twitter tweets," *Procedia Computer Science*, vol. 112, pp. 1964 – 1973, 2017, knowledge-Based and Intelligent Information and Engineering Systems: Proceedings of the 21st International Conference, KES-20176-8 September 2017, Marseille, France. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877050917314114

[13] O. Dreier, "Personality and the conduct of everyday life," *Nordic Psychology*, 2011.

[14] N. Majumder, S. Poria, A. Gelbukh, and E. Cambria, "Deep learning-based document modeling for personality detection from text," *IEEE Intelligent Systems*, vol. 32, no. 2, pp. 74–79, 2017.

[15] L. R. Goldberg, D. Sweeney, P. F. Merenda, and J. E. Hughes Jr, "Demographic variables and personality: The effects of gender, age, education, and ethnic/racial status on self-descriptions of personality attributes," *Personality and Individual differences*, vol. 24, no. 3, pp. 393–403, 1998.

[16] R. R. McCrae, "Personality theories for the 21st century," *Teaching of Psychology*, vol. 38, no. 3, pp. 209–214, 2011. [Online]. Available: https://doi.org/10.1177/0098628311411785

[17] L. A. Pervin and O. P. John, "Handbook of personality," *Theory and research*, vol. 2, 1999.

[18] H. E. Cattell and A. D. Mead, "The sixteen personality factor questionnaire (16pf)," *The SAGE handbook of personality theory and assessment*, vol. 2, pp. 135–178, 2008.

[19] H. J. Eysenck, *A model for personality*. Springer Science and Business Media, 2012.

[20] I. Briggs Myers and L. K. Kirby, "Introduction to type: a guide to understanding your results on the myers-briggs type indicator," *European English Version*, 2000.

[21] J. M. Digman, "Personality structure: Emergence of the five-factor model," *Annual review of psychology*, vol. 41, no. 1, pp. 417–440, 1990.

[22] M. S. Cole, H. S. Feild, and W. F. Giles, "Using recruiter assessments of applicants resume content to predict applicant mental ability and big five personality dimensions," *International Journal of Selection and Assessment*, vol. 11, no. 1, pp. 78–88, 2003.

[23] J. W. Pennebaker and L. A. King, "Linguistic styles: Language use as an individual difference." *Journal of personality and social psychology*, vol. 77, no. 6, p. 1296, 1999.

[24] J.-I. Biel, V. Tsiminaki, J. Dines, and D. Gatica-Perez, "Hi youtube!: personality impressions and verbal content in social video," in *Proceedings of the 15th ACM on International conference on multimodal interaction*. ACM, 2013, pp. 119–126.

[25] H. J. Escalante, H. Kaya, A. A. Salah, S. Escalera, Y. Gucluturk, U. Guclu, X. Baró, I. Guyon, J. J. Junior, M. Madadi *et al.*, "Explaining first impressions: modeling, recognizing, and explaining apparent personality from videos," *arXiv preprint arXiv:1802.00745*, 2018.

[26] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, 2010, pp. 94–101.

[27] R. M. Favaretto, P. Knob, S. R. Musse, F. Vilanova, and Â. B. Costa, "Detecting personality and emotion traits in crowds from video sequences," *Machine Vision and Applications*, vol. 30, no. 5, pp. 999–1012, 2019.

[28] F. Nihei, Y. I. Nakano, Y. Hayashi, H.-H. Huang, and S. Okada, "Predicting influential statements in group discussions using speech and head motion information," *Proceedings of the 16th International Conference on Multimodal Interaction*, 2014.

[29] D. Sanchez-Cortes, O. Aran, M. S. Mast, and D. Gatica-Perez, "A nonverbal behavior approach to identify emergent leaders in small groups," *IEEE Transactions on Multimedia*, vol. 14, no. 3, pp. 816–832, 2011.

[30] P. Senge, "The fifth discipline: The art and practice of organizational learning," *New York*, 1990.

[31] K. Kalimeri, B. Lepri, and F. Pianesi, "Going beyond traits: Multimodal classification of personality states in the wild," 12 2013, pp. 27–34.

[32] R. Jindal, M. Leekha, M. Manuja, and M. Goswami, "What makes a better companion? towards social and engaging peer learning," in *Proceedings of the Twenty-forth European Conference on Artificial Intelligence*. IOS Press, 2020.

[33] R. Morris, K. Kouddous, R. Kshirsagar, and S. Schueller, "Towards an artificially empathic conversational agent for mental health applications: System design and user perceptions," *Journal of Medical Internet Research*, 2018.

[34] L. R. Goldberg, "The development of markers for the big-five factor structure." *Psychological assessment*, vol. 4, no. 1, p. 26, 1992.

[35] C. M. Sims, "Do the big-five personality traits predict empathic listening and assertive communication?" *International Journal of Listening*, vol. 31, no. 3, pp. 163–188, 2017. [Online]. Available: https://doi.org/10.1080/10904018.2016.1202770

[36] C. O. Mawalim, S. Okada, Y. I. Nakano, and M. Unoki, "Multimodal bigfive personality trait analysis using communication skill indices and multiple discussion types dataset," in *Social Computing and Social Media. Design, Human Behavior and Analytics*, G. Meiselwitz, Ed. Cham: Springer International Publishing, 2019, pp. 370–383.

[37] R. Shah and R. Zimmermann, *Multimodal analysis of user-generated multimedia content*. Springer, 2017.

[38] H. W. Park, M. Gelsomini, J. J. Lee, and C. Breazeal, "Telling stories to robots: The effect of backchanneling on a child's storytelling," in *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI*. IEEE, 2017, pp. 100–108.

[39] R. Ishii, K. Otsuka, S. Kumano, and J. Yamato, "Predicting who will be the next speaker and when in multi-party meetings," *NTT Technical Review*, vol. 13, 07 2015.

[40] A. Kerry, R. Ellis, and S. Bull, "Conversational agents in e-learning," in *Applications and Innovations in Intelligent Systems XVI*, T. Allen, R. Ellis, and M. Petridis, Eds. London: Springer London, 2009, pp. 169–182.

International Journal of Multidisciplinary and Current Educational Research (IJMCER)
ISSN: 2581-7027 ||Volume|| 2 ||Issue|| 5 ||Pages|| 82-95||2020||

Watermarking of Multi-mode Counter Circuit using Hardware Efficient Algorithm

[1,]Jeebananda Panda, [2,]Divant Jain, [3,]Lavi Tanwar, [4,]Sunil Kumar
[1,2,3]*Department of Electronics and Communication Engineering, Delhi Technological University, Delhi 110042, India* [4]*Department of Electronics and Communication Engineering, Maharaja Agrasen Institute of Technology, Delhi 110086, India*

ABSTRACT: The paper proposes a property implant watermarking technique for the hardware designers to protect their Intellectual Property (IP) in digital integrated circuits. The technique introduces some extra states in the State Transition Graph of the digital circuit to embed the watermark. These extra states can further be used to prove the ownership in case of IP theft. The proposed scheme is intended for Finite State Machine (FSM) sequential circuit only due to the omnipresence of FSM. The efficiency of the proposed scheme can be seen in terms of less number of additional gates requirement and tough detection and removal of watermark from the original circuit. To demonstrate the process of watermarking, the proposed algorithm makes use of a Multipurpose Counter as a vehicle. The paper also presents the HDL simulations for different length of signature bits to ease the comparison between watermarked and non-watermarked circuits.

KEYWORDS: Property Implant technique, State Transition Graph, Intellectual Property, Finite State Machine, Signature sequence.

I. INTRODUCTION

With the emergence of newer technologies in semiconductor processing, the complexity in IC design and fabrication has increased. The whole process of hardware designing, starting from the designing to fabrication, is cumbersome and expensive task. From an economical point of view, today most of the ICs are tested and verified using computer aided design tools before going for the fabrication processes. In the whole design process there are possibilities that some of the important and critical file generated by the computer software may get stolen or got in the wrong hands. The designer must protect his design by securing his design in the first place itself. This task can be accomplished by embedding a unique code, or watermark, exploiting the IP's unique features. Fundamental requirements for a watermark scheme are that it be 1) Invisible, i.e., the watermark should be invisible to the end user, and it should not disturb the functionality of the original circuit, 2) Robust, i.e., the watermark is hard to remove and if infringed upon, it will destroy the circuit's intended functionality, and 3) Detectable, i.e., the watermark can be easily detected by the owner, which can be used to prove the ownership in case of IP theft. The signature sequence is only known to the owner.

A short overview of different watermarking schemes and their comparisons are presented in [1]. The paper also discusses major considerations and issues related with watermarking schemes, such as possible attacks on watermark, embedding cost, etc. One of the most popular schemes suited for IP cores are the constraint-based watermarks introduced in [2] and [3]. In these schemes the watermark is applied by defining additional design constraints which do not interfere with the functionality of the IP core, but will be utilised for embedding the watermark. The concept of watermarking has been proposed for many different levels of the hardware design process. The paper [4] and [5] discusses different algorithms which can be used to generate and detect watermarks at different abstraction levels. Paper [5], in particular presents the watermarking scheme at physical design level, where the watermark is embedded by using features characteristics of physical design level.

In [6] to [13], schemes have been proposed to embed the watermark at the top abstraction level of design i.e. at the behavioral level, where the designer describes the functionality of the design. The papers [6] and [7], presents a watermarking scheme for the sequential circuits, here the watermark is embedded by duplicating all the states in the State Transition Graph (STG) of the original circuit and adding new watermarking states. The modified STG thus obtained, contains watermarking states plus the duplicate states. This methodology, however robust and easily detectable, is not hardware efficient i.e. it would be impractical to use this scheme when there is large no. of states in the sequential circuit. In [8], the whole watermarking procedure, for the scheme presented in [6] and [7], has been illustrated through a sequence detector circuit. A modification to the scheme proposed in [6] and [7] is presented in [9] and [10]. Here the watermarking scheme is proved to be hardware efficient as compared to the original one.

It has been shown that in order to embed the watermark it is not necessary to duplicate all the original states and the scheme can be implemented using lesser no. of additional states. This reduces the hardware requirement manifolds. The paper shows the comparison between the modified scheme and the scheme using the results from [8] as a reference.The approaches in [11], [12] and [13], presents the watermarking scheme for finite state machine (FSM), here the watermark is implanted by exploiting some unutilized input vectors and modifying the states corresponding to these vectors to store the watermarking information. The watermark is thus detected by triggering a specific response with known input sequence.

The paper [14] presents a robust watermarking method for the audio signals using the invariance property of exponent moments technique to watermark the audio signals. The paper [15] presents a reversible image watermarking scheme in DWT domain, the scheme besides being robust, protects the high quality of the image from being tampered. A novel watermarking scheme for the digital contents using phase congruency techniques is presented in [16]. The papers [14-16] present IP protection of digital cover medium like audio and image. Even the design of a sequential machine can be an another form of digital cover medium or digital content whose IP also needs to be protected. In this paper, we will focus on the kind of watermarking scheme presented in [6] to [13] i.e. watermarking in the sequential circuits. We propose a modification to the watermarking scheme presented in [6] and [7]. The proposed scheme has been implemented with lesser no. of additional states required to embed the watermark and thus hardware efficient. Moreover, the scheme fulfils all the fundamentals requirements i.e. 1) Invisible, 2) Detectable and 3) Robust.The paper is organised in six sections. Section II presents a detailed description of the watermarking scheme proposed. Section III discusses the detection of presence of watermark in the circuit, followed by a brief discussion on robustness of the watermark. The implementation details of the watermarked circuit with different signature key bits are presented in section IV. Section V gives the experimental results obtained by simulating the Verilog HDL designs. Finally, section V concludes the paper.

II. WATERMARKING PROCEDURE

The watermark is embedded in the sequential circuit by introducing some extra states in the original state transition graph of the circuit. These states would be unique and will be used to detect the presence of watermark in the circuit.

For a given FSM let the set S = {s0, s1, . . .} represent the original states which describes the behaviour of the FSM. The set R = {r1, r2, . . .} is used to represent the watermarking states that will be used for detection of presence of the watermark in the circuit.The first step in watermarking process is to identify the secret key or signature for the given circuit. The key can be any arbitrary sequence of primary input combinations of the circuit. For example if there primary inputs, {x0, x1, x2}, then the sequence of input combinations can be represented as {y1, y2, y3, . . .} where {yi} represents a unique combination of inputs {x0, x1, x2}. But care should be taken not to take such a combination which is very common for the intended behaviour of the circuit i.e. it is advisable to take a combination which causes such states transitions which are rarely traversed as far as the normal functionality of the circuit is concerned. After the sequence input combination is finalised, we need to find the states which are traversed on applying that input combinations, let us denote these states by S' = {s'1, s'2, . . .}.

The steps to be followed to modify the state transition graph, in order to obtain a watermarked circuit are as follows:

a. Each state {ri} is formed by duplicating the state {s'i} and all its outgoing edges. Note that it is not necessary that all the states corresponding to the set {s'i} are different.

b. The outgoing edge from the starting state i.e. {s0} which corresponds to the transition {y1} is directed towards the state {r1}.

c. For the states {r1, r2, . . . , rn-1} the outgoing edge which corresponds to the transition {y2, y3, . . . , yn} is directed towards the states {r2, r3, . . . , rn}. Therefore every state {ri} has only one incoming edge, corresponding to the transitions {y1, y2, . . . ,yn}, and that also originating from the state {ri-1}.

d. All the outgoing edges from the states {r1, r2, . . . ,rn} which do not corresponds to the transitions {yi} are directed towards the original states {si}.

After the completion of the above procedure, we have a modified STG, which has watermark embedded in it. This modified sequential circuit follows all the fundamental requirements i.e. it is:

a. Invisible, meaning that the modified circuit has the same functionality as that of the original circuit.

b. Detectable, meaning that the watermark is easy to detect and prove the ownership in case of piracy.

c. Robust that is the watermark is hard to remove from the design.

In this paper we have used a multi-mode counter to demonstrate the watermarking procedure. A counter is a relatively simple sequential circuit and is easily available in the IC format by various vendors. Our multi-mode synchronous counter counts in three modes namely, 1) Binary, 2) Grey Code and 3) BCD Counter. It is a bidirectional counter i.e. it can count in both forward and backward directions. Further it has a synchronous reset and count enable pins, reset is used to set the counter in zero state and count enable is used to start or stop the counting. The STG of the counter is shown in figure 1 and the modified STG for an 18 bit watermark is shown in figure 2. Table 1 presents the relation between the present state, next state and the output for the circuit. We have shown the input as a three bit binary number, in which the LSB represents the "direction" input and the first two bits are for the "counter_type" i.e. the type of counting that the counter performs. The binary number {00} is assigned for binary counting, {01} is assigned for grey code counting and {10} is assigned for BCD counting. Further in figure 1, we have used colour scheme as – Blue, when output is '1'; Red, for grey code counting up; Magenta, for grey counting down; Black, for binary and BCD counting up and down. In figure 2, apart from these colours we have used colours – Green, for the transitions corresponding to the key or signature as input sequence i.e. {001, 011, 101, 010, 000, 100}; Orange, for the outgoing transitions from the watermarking states that do not result in a watermarking state as next state. The watermarking states are shown with a Yellow colour.

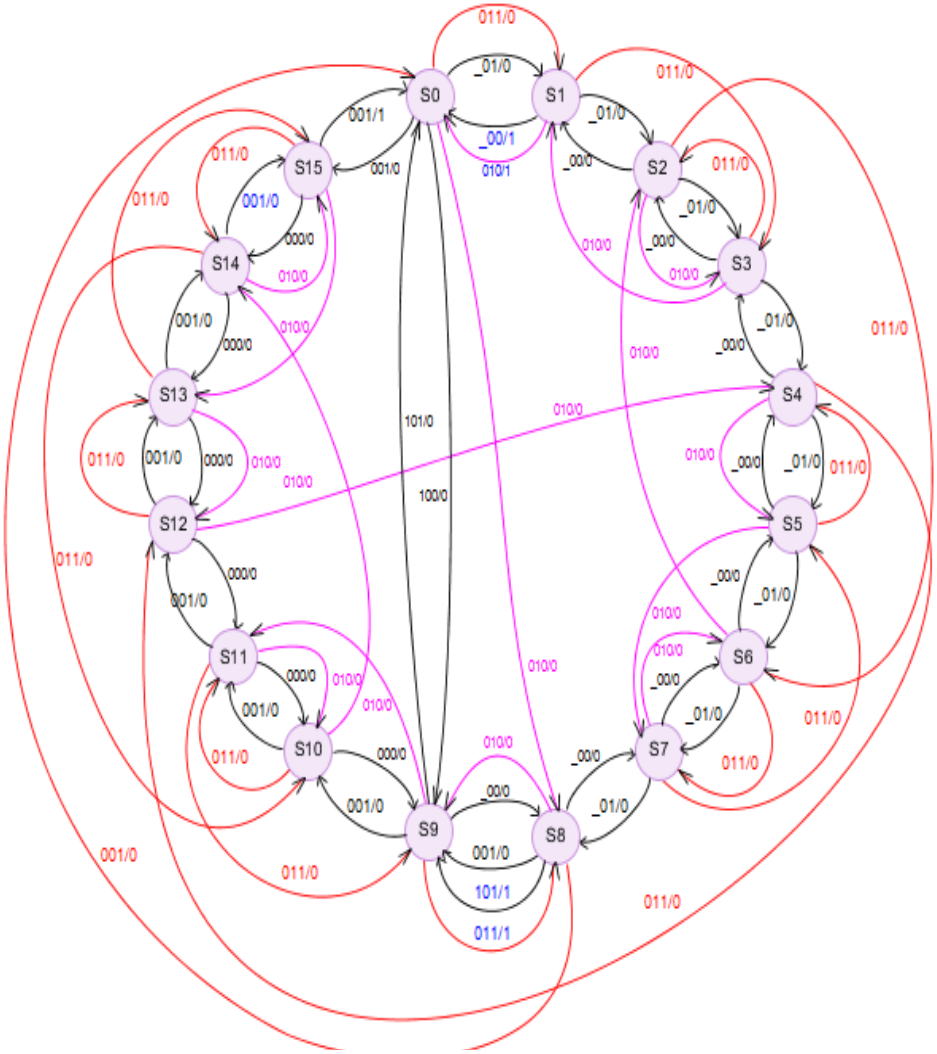


Figure 1. STG for the multipurpose counter circuit

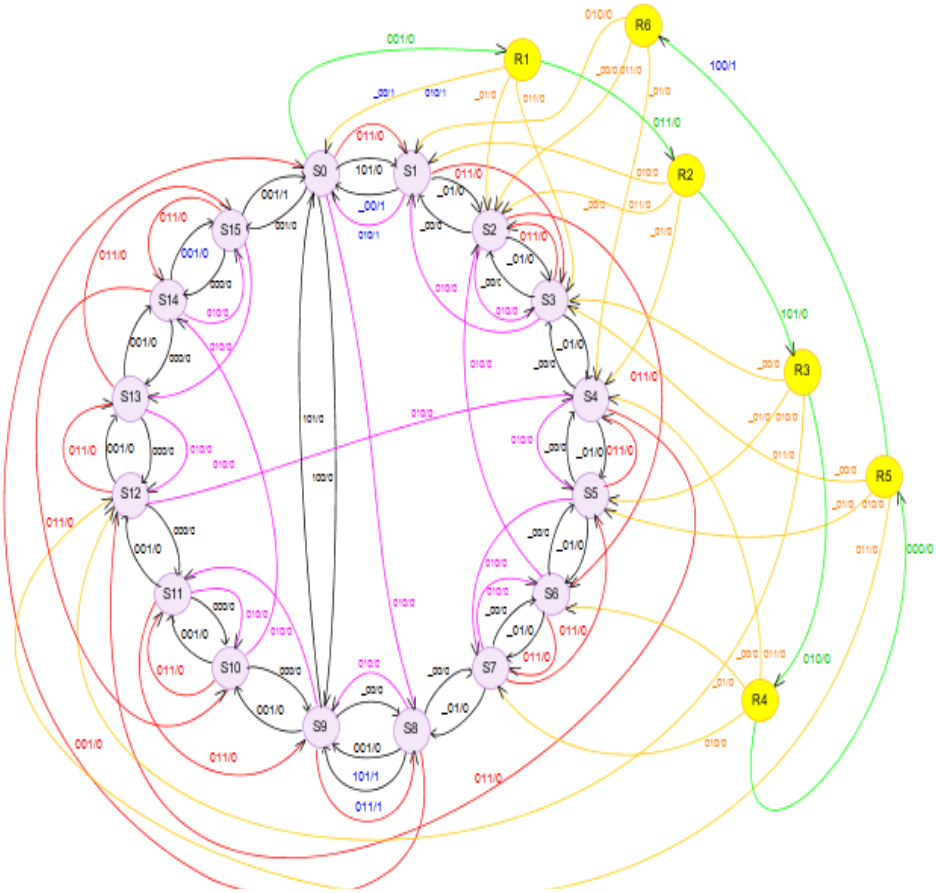


Figure 2. Modified STG for the multipurpose counter with 18-bit Watermark embedded in it

Table 1. State Table for the multipurpose counter

PRESENT STATE	NEXT STATE / OUTPUT						
	INPUTS						
	001	000	011	010	101	100	11_
S0	S1/0	S15/0	S1/0	S8/0	S1/0	S9/0	S0/0
S1	S2/0	S0/1	S3/0	S0/1	S2/0	S0/1	S0/0
S2	S3/0	S1/0	S6/0	S3/0	S3/0	S1/0	S0/0
S3	S4/0	S2/0	S2/0	S1/0	S4/0	S2/0	S0/0
S4	S5/0	S3/0	S12/0	S5/0	S5/0	S3/0	S0/0
S5	S6/0	S4/0	S4/0	S7/0	S6/0	S4/0	S0/0
S6	S7/0	S5/0	S7/0	S2/0	S7/0	S5/0	S0/0
S7	S8/0	S6/0	S5/0	S6/0	S8/0	S6/0	S0/0
S8	S9/0	S7/0	S0/0	S9/0	S9/1	S7/0	S0/0
S9	S10/0	S8/0	S8/1	S11/0	S0/0	S8/0	S0/0
S10	S11/0	S9/0	S11/0	S14/0	S0/0	S0/0	S0/0
S11	S12/0	S10/0	S9/0	S10/0	S0/0	S0/0	S0/0
S12	S13/0	S11/0	S13/0	S4/0	S0/0	S0/0	S0/0
S13	S14/0	S12/0	S15/0	S12/0	S0/0	S0/0	S0/0
S14	S15/1	S13/0	S10/0	S15/0	S0/0	S0/0	S0/0
S15	S0/0	S14/0	S14/0	S13/0	S0/0	S0/0	S0/0

As can be seen from the figure there are four watermarking states which are represented as {r1, r2, r3, r4, r5, r6}. There are two primary inputs. These two inputs make up three bits in total which are represented as {x0, x1, x2}, the first two bits represent the two bit input – "count_type" and the third bit is for the second input – "direction". Now the sequence of input combination or the signature or key chosen for the circuit shown in figure 2 is {y1, y2, y3, y4, y5, y6} = {001, 011, 101, 010, 000, 100}. On application of

this sequence of input the sequence of states traversed will be {s0, r1, r2, r3, r4, r5, r6}. Note that these states can only be traversed by this particular key and cannot be traversed by any other sequence of inputs.

III. DETECTION OF WATERMARK

The additional states, that are added to the STG of the original circuit, will be used for the detection of watermark embedded in the circuit and hence to prove the ownership in case of piracy. These additional states are thus also termed as watermarking states. From figure 2 it can be observed that states denoted by {r1, r2, r3, r4, r5, r6} represents the watermarking states. Further, it can be observed that these states have only one incoming edge and multiple outgoing edges. Moreover, every transition corresponding to the incoming edge in the watermarking states corresponds to the sequence of inputs {yi} i.e. the signature or the key. The above two properties can be stated as:

a. Every watermarking state can be reached only by its previous state i.e. the state {ri} can only be reached from the state {ri-1}.

b. Each and every watermarking state can only be traversed if and only if we apply the signature or key at the input correctly.

These two properties will be used to detect the presence of watermark. Note also that no other input sequence can result in transition through all the watermarking states. Thus only the secret key or the signature, which only the designer knows, can make the circuit transit through the watermarking states. Hence in case of piracy, the designer can use this key in the court-of-law and claim the ownership of the design.Figure 3 shows the simulation result of detection of watermark present in the modified circuit presented in figure 2. In this case, we have assigned binary value to the watermarking states as {r1 = 10000, r2 = 10001, r3 = 10010, r4 = 10011, r5 = 10100, r6 = 10101}. As can be seen from the figure 3, that on application of the input sequence as {001, 011, 101, 010, 000, 100}, which is the key chosen for this circuit, the sequence of states traversed are {s0, r1, r2, r3, r4, r5, r6} and the output will be 1 on the next rising edge of the clock after the state {r6} is reached, this is used only as an indication that the end of watermarking states has been reached. The states are represented in hexadecimal numbers in the figure as {00, 10, 11, 12, 13, 14, 15}. Similarly simulation for different keys sizes are shown in figure 4-8. For a 15 bit signature, the sequence of states traversed will be {s0, r1, r2, r3, r4, r5} on applying {001, 011, 101, 010, 000} as signature, this has been shown in figure 4. For a 12 bit signature, the sequence of states traversed will be {s0, r1, r2, r3, r4} on applying {001, 011, 101, 010} as signature, this has been shown in figure 5. Similarly for 9 and 6 bit signatures the sequence of states traversed will be {s0, r1, r2, r3} and {s0, r1, r2} on applying {001, 011, 101} and {001, 011} as signatures respectively, this has been depicted in figures 6 and 7. It should be noted that the robustness of the circuit decreases with decreasing size of the key.

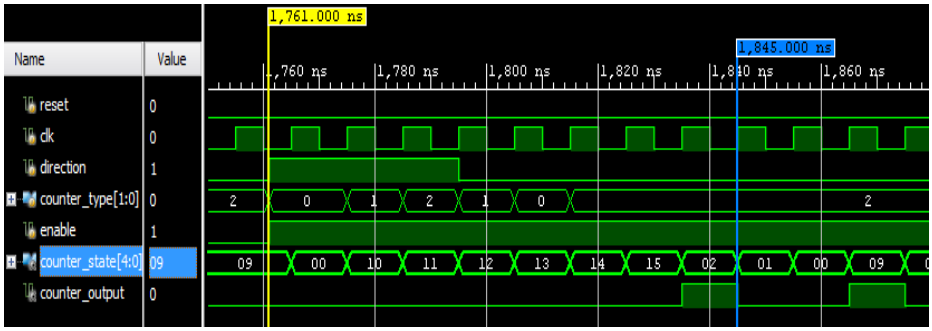


Figure 3. Detection of Watermark (18-bit watermarked circuit)

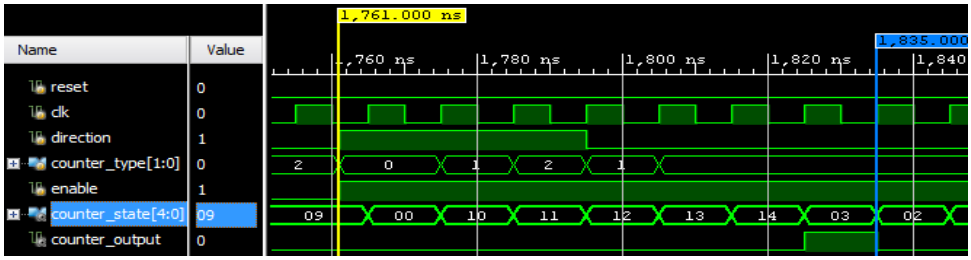


Figure 4. Detection of Watermark (15-bit watermarked circuit)

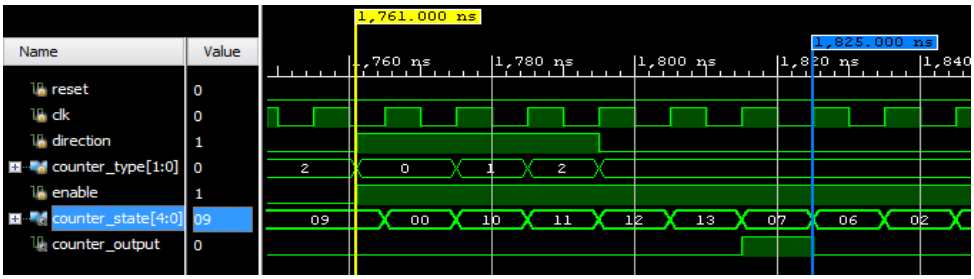


Figure 5. Detection of Watermark (12-bit watermarked circuit)

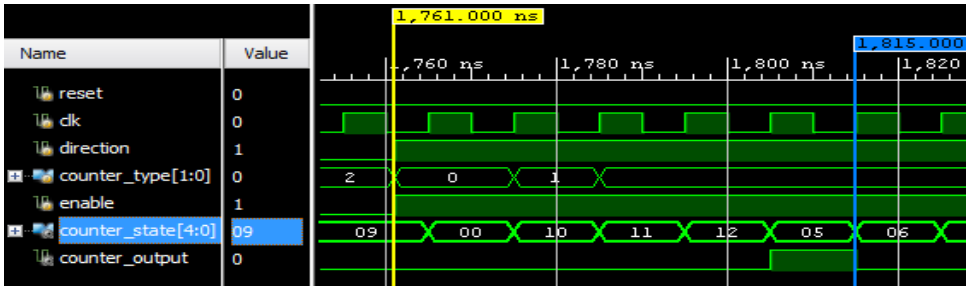


Figure 6. Detection of Watermark (9-bit watermarked circuit)

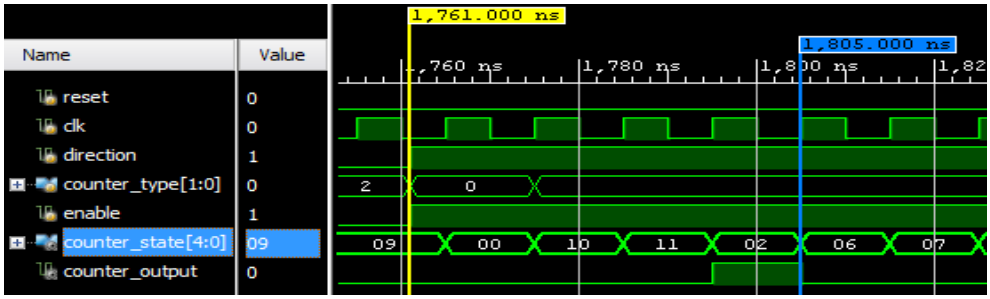


Figure 7. Detection of Watermark (6-bit watermarked circuit)

Thus, the watermark is easy to detect and at the same time is robust i.e. if it is infringed upon, that will result in damage to the circuit and its functionality. The watermarking states form an integral part of the circuit and are necessary for the normal functioning of the circuit. This property can be verified from the simulations (presented in section V) of the watermarked circuit for various operations. Here the difference between the state transition of the watermarked and non-watermarked circuits should be noted. Thus, due to combinatorial complexity, it is not feasible to erase or delete the watermark from the FSM with finite resources and time. Therefore, the presented watermark scheme is robust to attacks and easy to apply.

IV. IMPLEMENTATION OF WATERMARK

To illustrate the formulation of an efficient watermarking signature sequence, we have used multi-mode counter as the base sequential circuit. The modified STG of the counter circuit for five different sizes of keys (that varies from 6 to 18 bits) were simulated and analysed in order to formulate a most efficient signature sequence with minimal hardware overweight and as efficient as un-watermarked circuit in terms of timing parameters. Although the security of the circuit increases with the size of the key, but care should be taken that the watermark is embedded with minimal hardware overhead. Thus there is a trade-off between the key size and hardware overhead associated with it. Figures 8-11 and figure 2, shows the modified STG for different sizes of keys. Hardware requirements and timing specifications for each of the modified circuit is given in section V. For an 18 bit watermark, we have selected the key as input sequence {001, 011, 101, 010, 000, 100}. This input sequence is a rare set of input combination for anyone to imagine and hence can be used as a signature to watermark the design. On the other hand, for a 6 bit watermark the signature sequence selected is {001, 011}, although it is not a common operation, but it is less secure as compared to the 18 bit watermark because of less no. input sequence. Thus, as far as the security of the watermark is concerned, the watermark with more no. of input bits is more secure.

As can be seen from the figure 2 and 8-11 with increasing signature sizes the number of watermarking states, shown in yellow, increases and this will increase the overall hardware requirement for the circuit.

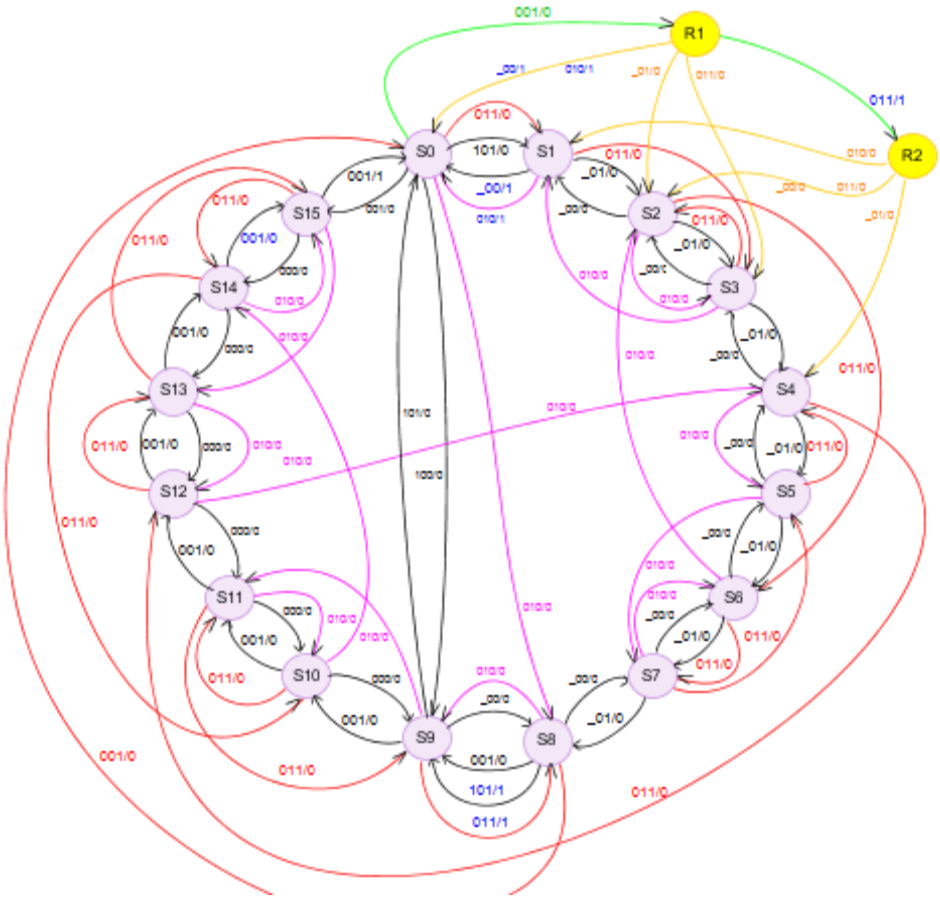


Figure 8. Modified STG for the multipurpose counter with 6-bit Watermark embedded in it

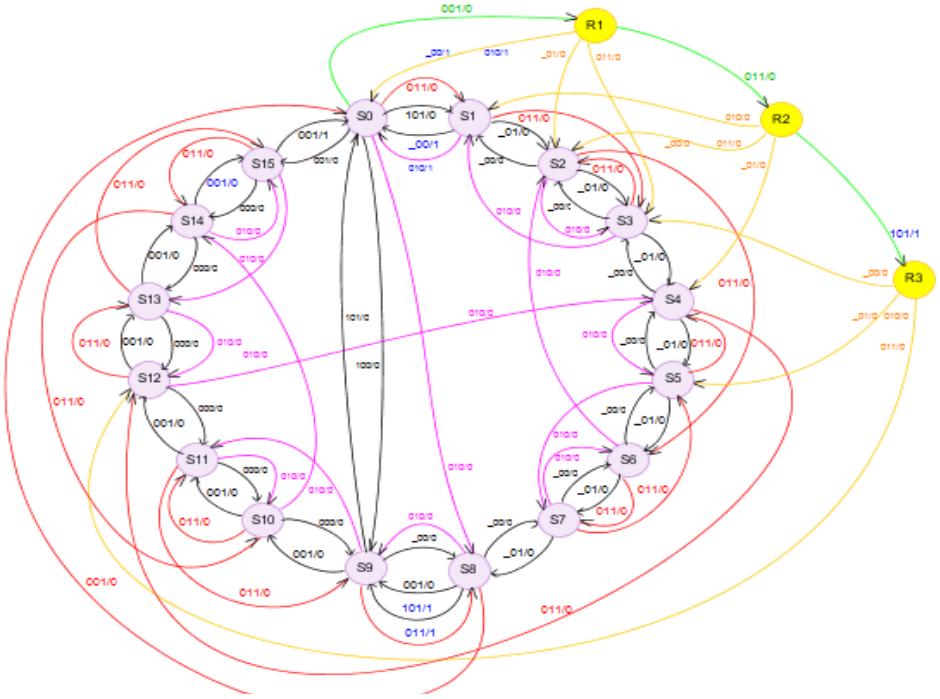


Figure 9. Modified STG for the multipurpose counter with 9-bit Watermark embedded in it

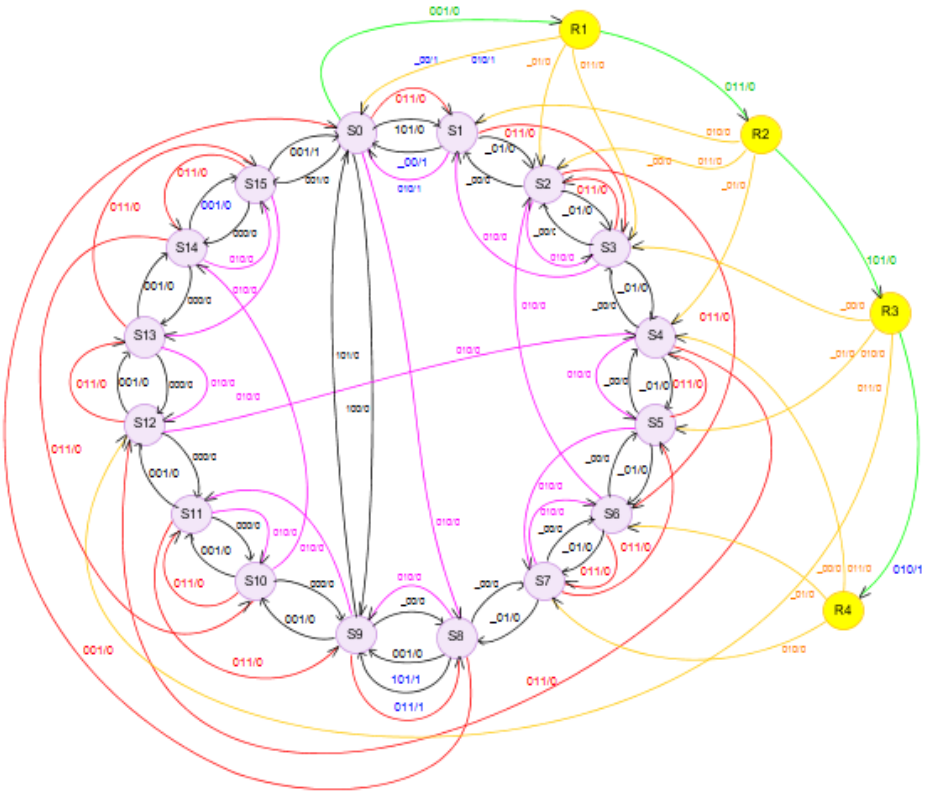


Figure 10. Modified STG for the multipurpose counter with 12-bit Watermark embedded in it

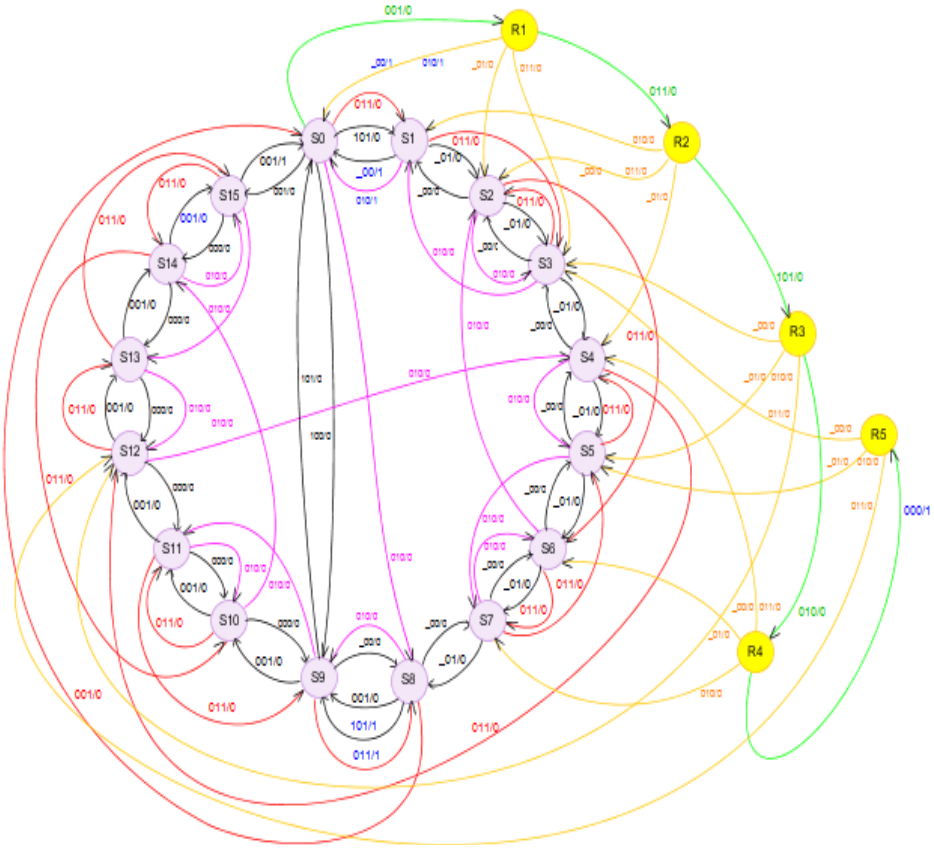


Figure 11. Modified STG for the multipurpose counter with 15-bit Watermark embedded in it

V. EXPERIMENTAL RESULTS

The design of the multi-mode counter was implemented using the Verilog HDL. The Xilinx Vivado Suit software was used synthesize and simulate the design. Two separate modules for the watermarked and non-watermarked circuit were made. The simulation result for the three modes of counting, shown in figure 12-17, are for the original circuit i.e. non-watermarked circuit. Figure 18-23 shows the simulation results for the watermarked circuit. It can be observed that simulations for original circuit exactly match the simulations for the watermarked circuit as far as the functionality of the circuit is concerned. It is an important result that verifies the invisibility of the watermark i.e. the end user should not be aware that a watermark is embedded in the circuit. However, it should be noted that the states that are traversed on applying the same input to both the circuits vary greatly.

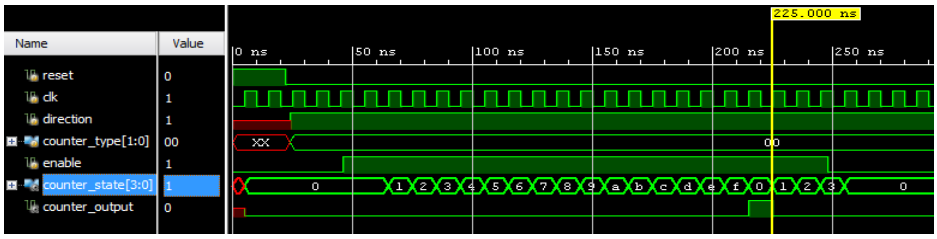


Figure 12. Binary counting up (original circuit)

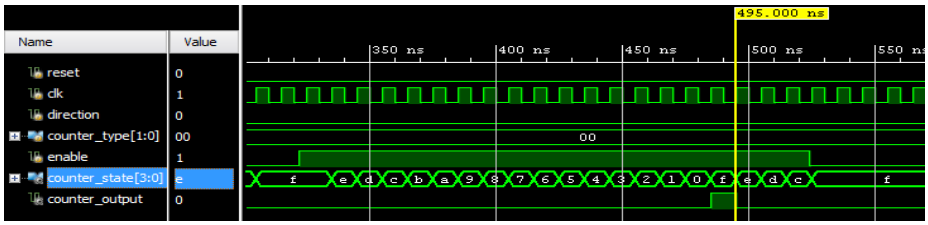


Figure 13. Binary counting down (original circuit)

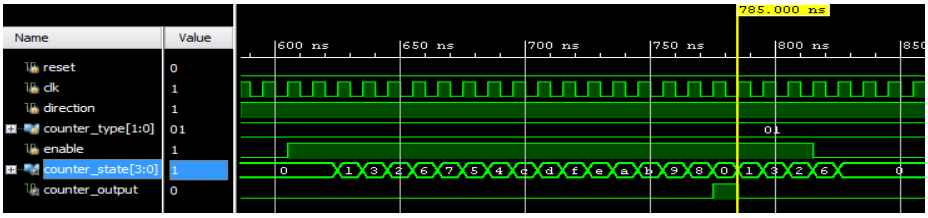


Figure 14. Gray code counting up (original circuit)

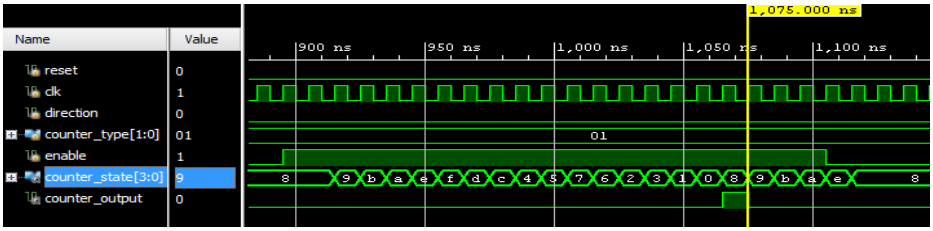


Figure 15. Gray code counting down (original circuit)

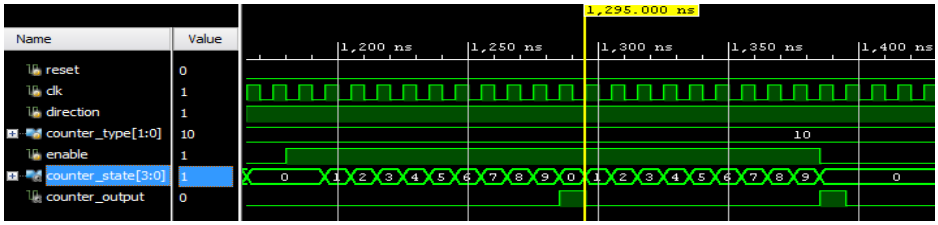


Figure 16. BCD counting up (original circuit)

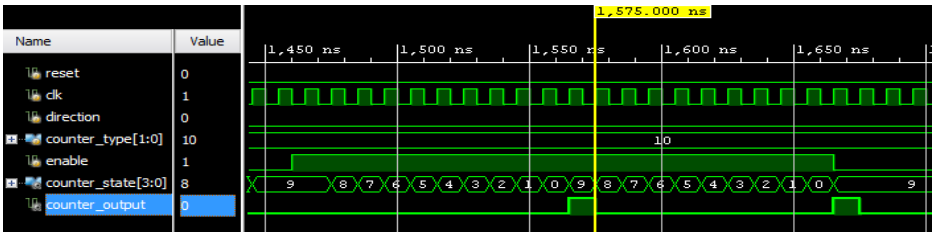


Figure 17. BCD counting down (original circuit)

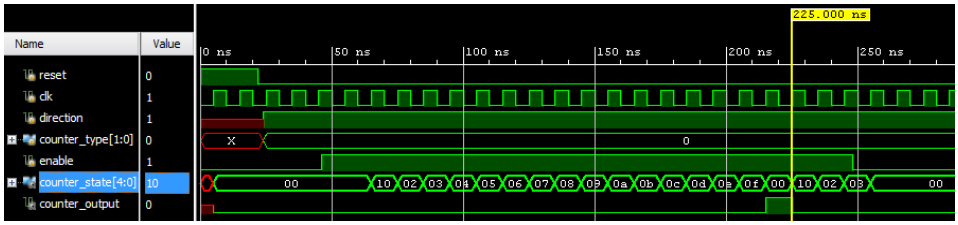


Figure 18. Binary counting up (watermarked circuit)

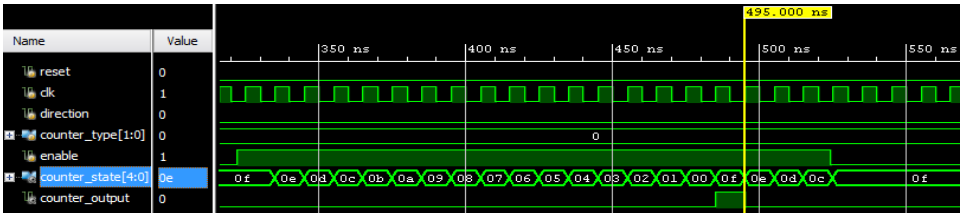


Figure 19. Binary counting down (watermarked circuit)

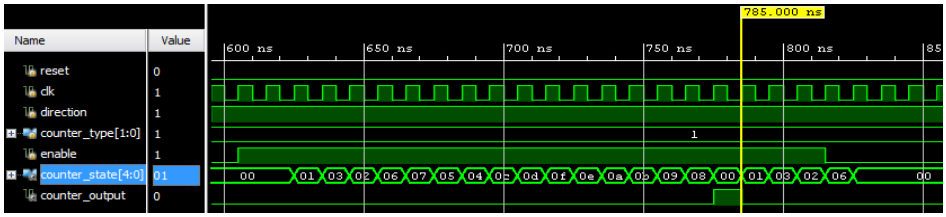


Figure 20. Gray code counting up (watermarked circuit)

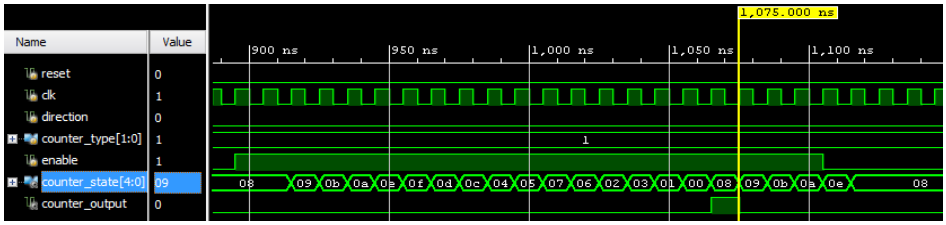


Figure 21. Gray code counting down (watermarked circuit)

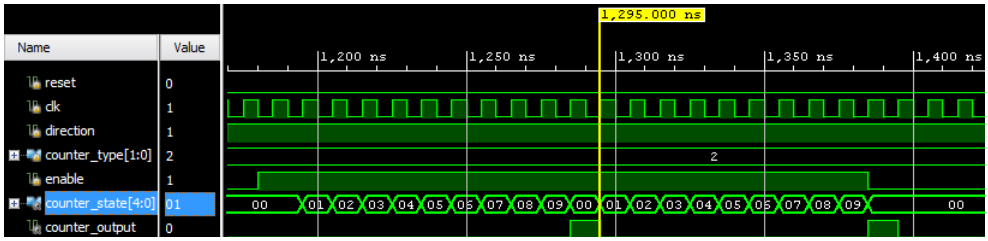


Figure 22. BCD counting up (watermarked circuit)

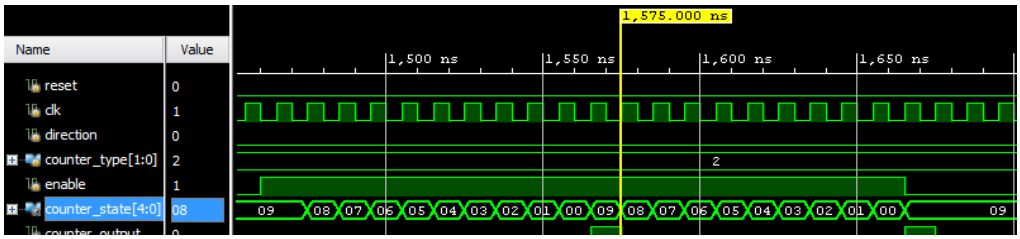


Figure 23. BCD counting down (watermarked circuit)

The hardware requirements and timing specifications for the original design and the watermarked design with different key sizes are given in table 2 and table 3. It can be seen that the modified circuit requires only few extra hardware components as compared to the original circuit. Further, it should be emphasized on the fact that among the different watermarked designs, the best would be to use the one with 12 bit watermarking key. From table 2, it can be observed that among the five available signatures the 12 bit and 15 bit watermarked circuits can be used efficiently. For others signatures, either the lengths are too short (6 and 9 bit) or the signature hardware requirements are more (15 bit). From table 3, we see that among the 12 bit and 15 bit signature, the 12 bit signature is more close to the original design parameters. Thus, the watermarked design with 12 bit signature has its parameters more close to the original design and also 12 bit signature length is large enough that it can be efficiently used to watermark the circuit.

Table 2. Hardware Requirements

Functional Unit	Original Design	Watermarked design				
		6 bit	9 bit	12 bit	15 bit	18 bit
Flip Flop & Latches	7	8	8	8	8	8
IO Buffers	12	12	12	12	12	12
Clock Buffers	1	1	1	1	1	1
Multiplexers	1	0	1	0	2	4
LUTs	27	40	41	42	40	50

Table 3. Timing Specifications

Timing Specification	Original Design	Watermarked design				
		6 bit	9 bit	12 bit	15 bit	18 bit
Maximum Frequency (MHz)	269.26	262.64	268.40	265.91	263.07	265.51
Minimum input arrival time before clock (ns)	5.088	5.143	5.722	5.135	5.388	5.333
Maximum output required time after clock (ns)	4.089	4.131	4.110	4.131	5.731	5.786

It should be noted that if the counter circuit presented in this paper would have been watermarked using the scheme presented in [6] and [7], we would require double the original states plus the watermarking states as the total number of states for the watermarked circuit, which would be 36 states. The proposed method, on the other hand, requires only the watermarking states as the additional states i.e. a total of 20 states for the watermarked STG. Thus the proposed scheme requires much less states and therefore can be implemented with minimal hardware overhead.

VI. CONCLUSION

The Verilog HDL simulation for all the five modified STGs with different length of signature sequence proves the functional similarity of the watermarked and non-watermarked circuits, which fulfils the foremost condition of any watermarking scheme. It also has been shown that the watermark can be easily applied and detected to prove the ownership in case of any kind of piracy of the original design.

Moreover, the watermarking scheme is robust and if it is infringed upon, it will destroy the intended behaviour of the design and hence the whole circuit. Finally, it has been shown that the signature key should be selected in such a way that it minimise the hardware overhead caused and at the same time rare and large enough that it is not possible for anyone, except the owner, to access the key and use it to prove the authorship of the design. Base on the simulation results, it is observed that the modified STG with 12 bit signature sequence needs very few additional hardware overhead with very close timing parameters to that of un-watermarked circuit.As a future scope, this type of analysis can be applied to any FSM to obtain an efficient watermark signature sequence, which can be used to watermark the original FSM in order to obtain a hardware efficient watermarked FSM by property implant scheme. The similar technique can also be adopted for watermarking of commonly used digital blocks in digital system design.

REFERENCES

[1] A.T.A. Hamid, S. TaharE.M.Aboulhamid,"IP Watermarking Techniques: Survey and Comparison", 3rd IEEE International Workshop on System-on-Chip for Real-Time Applications, IEEE, Canada,**(2003)**, pp. 60–65.

[2] A.B. Kahng, J. Lach,W.H.M. Smith, S.Mantik,L.L. Markov,M. Potkonjak, P. Tucker, H. Wang, G. Wolfe, "Watermarking Techniques for Intellectual Property Protection", 35th Design Automation Conf., IEEE, USA,**(1998)**,pp. 776–781.

[3] D. Kirovski, Y.Y. Hwang,M. Potkonjak,J. Cong, J,"Intellectual Property Protection by Watermarking Combinational Logic Synthesis Solutions", IEEE/ACM Int. Conf. on Computer Aided Design, IEEE, USA,**(1998)**,pp. 194–198.

[4] E. Charbon, "Hierarchical Watermarking in IC Design", Proceedings of IEEE Custom Integrated Circuit Conf., IEEE, USA,**(1998)**, pp. 295–298.

[5] E. Charbon, I. Torunoglu,"Watermarking Layout Topologies",Proceedings of IEEE Asia and South-Pacific Design Automation Conf., IEEE, Hong Kong,**(1999)**, pp. 213–216.

[6] A.L. Oliveira,"Techniques for the Creation of Digital Watermarks in Sequential Circuit Design", IEEE Transactions on Computer-aided Design of Integrated Circuits and Systems, vol. 20, no. 9,**(2001)**, pp. 1101–1117.

[7] A.L. Oliveira,"Robust Techniques for Watermarking Sequential Circuit Designs", Proceedings of 36th IEEE Design Automation Conf., IEEE, USA,**(1999)**, pp. 837–842.

[8] S. Subbaraman, P.S. Nandgawe,"Intellectual Property Protection of Sequential Circuits Using Digital Watermarking", 1st Int. Conf. on Industrial and Information Systems, IEEE, Sri Lanka,**(2006)**, pp. 556–560.

[9] J. Panda, S. Malik, N. Pandey, A. Bhattacharyya,"A Modified Hardware Efficient Watermarking Scheme for Intellectual Property Protection in Sequential Circuits", Int. Journal of Electronics Communication and Computer Engineering, vol. 5, no.4, **(2014)**, pp. 741– 746.

[10] J. Panda, A. Bharadwaj, N. Pandey, A. Bhattacharyya,"Hardware efficient watermarking technique for finite state sequential circuit using STG", Int. Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering, vol. 2, no. 7, **(2014)**, pp.1670–1674.

[11] I. Torunoglu, E. Charbon,"Watermarking-Based Copyright Protection of Sequential Functions", IEEE Journal of Solid-State Circuits,vol. 35, no. 3, **(2000)**, pp. 434-440.

[12] A. Cui, C.H. Chang, S. Tahar, A.T.A. Hamid, "A Robust FSM Watermarking Scheme for IP Protection of Sequential Circuit Design", IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 30, no. 5, **(2011)**, pp. 678–690.

[13] M. Lewandowski, R. Meana, M. Morrison, S. Katkoori, "A Novel Method for Watermarking Sequential Circuits", IEEE International Symposium on Hardware-Oriented Security and Trust, IEEE, USA,**(2012)**, pp. 21–24.

[14] X.Y. Wang, Q.L. Shi, S.M. Wang,H.Y. Yang, "A blind robust digital watermarking using invariant exponent moments", International Journal of Electronics and Communications, vol. 70, no. 4, **(2016)**, pp. 416-426.

[15] T.S. Nguyen, C.C. Chang, X.Q. Yang,"A reversible image authentication scheme based on fragile watermarking in discrete wavelet transform domain", International Journal of Electronics and Communications, vol. 70, no. 8,**(2016)**, pp. 1055–1061.

[16] M.R. Nayak, J. Bag, S. Sarkar, S.K. Sarkar, "Hardware implementation of a novel water marking algorithm based on phase congruency and singular value decomposition technique", International Journal of Electronics and Communications, vol. 71, **(2017)**,pp. 1–8.

Authors



Jeebananda Panda born on 15th Feb, 1968 in Odisha, India. He got graduated in Electrical engineering and subsequently in Electronics and Communication Engineering in 1988 and 1989 respectively. He got his M.E degree in Applied Electronics specialization from Bharathiyar University in 1992 and got his Ph.D. degree from University of Delhi. Presently working as Professor in the Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India. His field of research is Watermarking of Digital Data.



Divant Jain born in Delhi, India, in 1995. He did hisB.Tech in Electronics and Communication Engineering from Delhi Technological University, Delhi, India. His areas of interests are in RISC processors and computer architecture.



Lavi Tanwar born in Delhi, India, in 1992.Presently working as an Assistant Professor in the Department of Electronics & Communication Engineering,Delhi Technological University, Delhi, India and also pursuing PhD in the field of Digital Watermarking from Delhi Technological University. She received 5-year Integrated B.Tech (Electronics and Communication Engineering) + M.Tech (Intelligent Systems and Robotics) degree from Gautam Buddha University, Greater Noida, Uttar Pradesh, India in 2015.



Sunil Kumar born on 10th March 1970. He graduated in Electronics and Communication Engineering from Nagpur University. He got his M.Tech in ECE from Delhi University and Ph.D from Singhania University. Presently working as Professor and Head of the Department of Electronics and Communication Engineering, Maharaja Agrasen Institute of Technology, Delhi, India. His field of research are image processing and digital watermarking.