SCHOLARLY PUBLICATIONS

A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT

@DTU

(104th Edition)

AUGUST 2021

BY: CENTRAL LIBRARY

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

PREFACE

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

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

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

This issue contains the information published during **August, 2021**. The arrangement of the contents is alphabetical. The full text of the article which is either subscribed by the university or available in the web is provided in this bulletin.

Central Library

CONTENTS

1. A Comprehensive Review of Adversarial Learningand Impact of Unsharing Weights Across Classes, ***8.S Vignesh, Bhawna***, ***8.Shubham Anand*** and ***3.Dr. Shailender Kumar***, CSE, DTU

2. A deep multimodal network based on bottleneck layerfeatures fusion for action recognition, ***6.Tej Singh*** and ***3.Dinesh Kumar Vishwakarma***, IT, DTU

3. A study on transformer-based Object Detection, ***8.Hritik Vaidwan***, ***8.Nikhil Seth***, ***3.Anil Singh Parihar*** and ***3.Kavinder Singh***, Physics and CSE, DTU

4. A Survey on Intrinsically MotivatedReinforcement Learning, ***3.Shweta Meena***, ***8.Abhay Singh Hyanki*** and ***8.Tarun Kumar***, Software Engineering, DTU

5. AE-DCNN: Autoencoder Enhanced DeepConvolutional Neural Network For MalwareClassification, ***8.Shashank Kumar***, ***8.Shivangi Meena***, ***8.Savya Khosla*** and ***3.Anil Singh Parihar***, CSE, DTU

6. An Analysis on Abrupt Fall of Indian Stock Markets, ***6.Sadhvi Mehra*** and ***3.Sumedha Seniaray***, Applied Mathematics, DTU

7. An efficient adsorbent for the removal of Zn2+Cd2+and Hg2+from the real industrial effluents, ***6.D. Chauhan***, ***3.A. Kumar*** and ***3.S. G. Warkar***, Chemistry, DTU

8. Analog Analysis of Novel Ferroelectric-DualMaterial Oxide Stack-Double Gate FET, ***6.Isha*** and***3.Rishu Chaujar***, Physics, DTU

9. Analysis of Automated text generation using Deeplearning, ***3.Dr. Manoj Kumar***, ***8.Abhishek Singh***, ***8.Arnav Kumar*** and ***8.Ankit Kumar***, CSE, DTU

10. Analysis of Time Series ForecastingTechniques for Indian Automotive Industry, ***8.PIYUSH JAIN*** and ***8.PARV ARORA***, Production, DTU

11. Antimicrobial Finishing of Cotton Textiles using Silver Intercalated Clay, Rosina Begam, Mangala Joshi and ***3.Roli Purwar***, Chemistry, DTU

12. Application of Hybrid Metaheuristic Technique to Study Influenceof Core Material and Core Trench on Performance of Surface InsetPMSM, ***3.Monika Verma***, ***3.Mini Sreejeth*** and ***3.Madhusudan Singh***, Electrical, DTU

13. Assessing the impact of negative WOM on diffusion process, ***3.Deepti Aggrawal***, Mohini Agarwal, Rubina Mittal and Adarsh Anand, USME, DTU

14. Attention U-Net Architecture for Dental ImageSegmentation, ***8.Pratyush Harsh***, ***8.Ratul Chakraborty***, ***8.Shaurya Tripathi*** and ***3.Dr. Kapil Sharma***, IT, DTU

15. Comparative analysis of acquisition methods in digital forensics, Himanshu, Shobha Bhatt and ***3.Geetanjali Garg***, CSE, DTU

16. Conceptual Design of an Unmanned Aerial Vehicle forMars Exploration, ***8.Ishan Mishra***, ***8.Aayush Kumar*** and ***8.Vanshaj Malhotra***, Mechanical, DTU

17. Convolutional Neural Network for Covid-19 detection from X-ray images, Anju Jain, Saroj Ratnoo and ***3.Dinesh Kumar***, Electroncis, DTU

18. COVID-19 Detection from Chest X-Rays and CT Scans using Dilated Convolutional Neural Networks, ***3.Shailender Kumar***, ***8.Kumar Abhishek***, ***8.Kritika*** and ***8.Kashish Singh***, CSE, DTU

19. COVID-19 Tweet Analysis using Hybrid Keyword Extraction Approach, ***8.Sakshi Vatsa***, ***8.Surbhi Mathur***, ***8.Mansi Garg*** and***3.Rajni Jindal***, CSE, DTU

20. CRISPR/Cas9 gene editing: New hope for Alzheimer's diseasetherapeutics, ***6.Shanu Bhardwaj***, Kavindra Kumar Kesari, Mahesh Rachamalla, Shalini Mani, Ghulam Md. Ashraf,Saurabh Kumar Jha, Pravir Kumar, Rashmi K. Ambasta, Harish Dureja, Hari Prasad Devkota,Gaurav Gupta, Dinesh Kumar Chellappan, Sachin Kumar Singhm, Kamal Dua, Janne Ruokolainen,Mohammad Amjad Kamal, Shreesh Ojha and Niraj Kumar Jha, Biotechnology, DTU

21. Customer Relationship Management of Different Companies, ***8.Mozammil Hassan*** and ***7.Sadique Hassan***, Mechanical and Environmental, DTU

22. DeepABM: SCALABLE AND EFFICIENT AGENT-BASED SIMULATIONS VIA GEOMETRICLEARNING FRAMEWORKS - A CASE STUDY FOR COVID-19 SPREAD ANDINTERVENTIONS, ***8.1.Ayush Chopra***, Ramesh Raskar, Esma S. Gel, CSE, DTU

23. Defining the Undefined, Aman Chaudhary and ***6.Luckshay Batra***, DTU

24. DESIGN AND ANALYSIS OF RING MULTIBANDMETAMATERIAL, ABSORBER ***3.Priyanka Jain*** and ***8.Keshav Bhati***, Electroncis, DTU

25. Design and Development of SAW Sensorbased Wireless Target type Flow Transmitter, Praveen Maurya, ***7.1.Sayyed Faizan Ali*** and Nirupama Mandal, Electroncis, DTU

26. Design and Performance Analysis for UndergroundTrains using Optisystem, ***3.Anurag Chauhan***, ***8.Rishabh Verma***, ***8.Ashish Joshi*** and ***8.Rishabh Sharma***, Electronics, DTU

27. Design, synthesis and self-assembly of amide-linked dendron-basednon-ionic amphiphiles, Ashwani Kumar, ***3.Ram Singh*** & Yogesh Kumar Tyagi, Chemistry, DTU

28. Designing of Serial Asynchronous to AnalogResolver Format Converter, ***7.Vijendra Kumar Singh*** and ***3.Dr. Yashna Sharma***, Electronics, DTU

29. Drying Kinetics and Economic Analysis of Bitter 1 Gourd Flakes Drying2 inside Hybrid Greenhouse Dryer, Asim Ahmad, Om Prakasha and ***3.Anil Kumar***, Mechanical, DTU

30. Dynamic Visualization and Analysis of Government Responses-A Support System to Control Pandemic Situations, Neelu Chaudhary, Hardeo Kumar Thakur, and ***6.1.Rinky Dwivedi***, CSE, DTU

31. Edge Detection in Digital Images Using Guided L0SmoothenFilter and Fuzzy Logic, Akshi Kumar and ***6.Sahil Raheja***, IT, DTU

32. Efficient Methods to Improve the Performance ofSupervised Learning Models, ***8.Mohit Kumar***, ***8.Pragya Yadav***, ***8.Harman Singh*** and ***3.Ankita Arora***, Electrical, DTU

33. Emission of respirable dust from stone quarrying, potential healtheffects, and its management, ***6.Saurav Kumar Ambastha*** & ***3.Anil Kumar Haritash***, Envt, DTU

34. Emissions from homogeneous charge compression ignition (HCCI)engine using different fuels: a review, ***8.Sanjeev Kumar Verma***, ***8.Subhashish Gaur***, ***8.Tabish Akram***, Samsher Gautam and ***3.Anil Kumar***, Mechanical, Production & Industrial and Environment, DTU

35. Evaluation of the residual stress of HVOF sprayed carbon coatingafter wear testing conditions using ANN coupled Taguchi approach, ***3.Ankit Tyagi***, ***3.Qasim Murtaza*** and RS Walia, Mechanical, DTU

36. Experimental investigation and artificial neural network modelingof performance and emission of a CI engine using orange peel oildieselblends, ***7.Sandeep Kumar Barnawal*** and ***3.Naveen Kumar***, Mechanical, DTU

37. Experimental Studies on Frozen Soil, ***7.Kshitij Gaur*** and ***3.Anil Kumar Sahu***, DTU

38. Experimental, Computational and Mathematical Analysis of HybridAbrasive Flow Machining Process, ***6.Sachin Dhull***, ***3.R. S. Walia***, ***3.Qasim Murtaza*** and ***3.M. S. Niranjan***, Mechanical, DTU

39. Extracting Notes From Youtube Video, ***3.Manoj Kumar***, ***8.Sumit Gaurav***, ***8.Suhail Akhtar*** and ***8.Siddharth Varshney***, CSE, DTU

40. Feasibility Study on Production of Hydrogen from Renewable Sources, ***8.Saurav Yadav***, ***8.Saurabh Agase*** and ***3.Dr. J.P. Kesari***, Mechanical, DTU

41. Feature Enhanced Capsule Networks for RobustAutomatic Essay Scoring, ***8.Arushi Sharma***, ***8.Anubha Kabra*** and ***3.Rajiv Kapoor***, DTU

42. FOXO and related transcription factors binding elements in the regulationof neurodegenerative disorders, ***8.Vaibhav Oli***, ***6.Rohan Gupta*** and ***3.Pravir Kumar***, Biotechnology, DTU

43. FPGA bases design of multifunction ALU, ***3.Anup Kumar***, ***8.Manish Kumar***, ***8.Gouri Shankar Jha*** and ***3.Kriti Suneja***, Electronics, DTU

44. Free Space Optical Wireless Communicationusing Lasers Array for Data Transmission, ***8.Durgesh***, ***8.Deepak Ghanshala***, ***8.Nachiketa Panwar*** and ***3.Piyush Tewari***, Electronics, DTU

45. Functional connectivity between frontal/parietal regionsand MTL–basal ganglia during feedback learningand declarative memory retrieval, ***6.SWATI AGRAWAL*** and ***3.RINKU SHARMA*** and VIJAYAKUMAR CHINNADURAI, PHYSICS, DTU

46. GRADE: A GRAPH BASED DATA-DRIVEN SOLVER FORTIME-DEPENDENT NONLINEAR PARTIAL DIFFERENTIALEQUATIONS, ***8.Yash Kumar*** and Souvik Chakraborty, Mechanical, DTU

47. Gratitude resentment and appreciation scale (GRAT-16): analyzingpsychometrics properties in the Indian context, ***3.Naval Garg*** and Mehak, USME, DTU

48. Implementation of High Performance 4-Bit ALU usingDual Mode Pass Transistor Logic, ***3.Anurag Chauhan***, ***8.Kirandeep Kaur Saini***, ***8.Nitin Rajput*** and ***8.Rushil Domah***, Electronics, DTU

49. Intelligent Algorithm based Efficient Planning of Electric Vehicle Charging Station: A Case Study of Metropolitan City of India, ***6.M. Bilal*** and ***3.Mohammad Rizwan***, Electrical, DTU

50. Intelligent Chatbot Using GNMT, SEQ-2-SEQTechniques, ***3.Rajesh Kumar Yadav***, ***8.Sahil Harwani***, ***8.Satyendra Kumar Maurya*** and ***8.Sachin Kumar***, CSE, DTU

51. Investigation of a combined refrigeration and air conditioning system based on two-phase ejector driven by exhaust gases of natural gas fueled homogeneous charge compression ignition engine, Abdul Khaliq, Bandar A. Almohammadi, Mathkar A. Alharthi, ***6.Mohd Asjad Siddiqui*** and ***3.Rajesh Kumar***, Mechanical, DTU

52. Measuring landslide vulnerabilitystatus of Chukha, Bhutan usingdeep learning algorithms, Sunil Saha, ***3.Raju Sarkar***, Jagabandhu Roy, Tusar Kanti Hembram, Saroj Acharya,Gautam Thapa and Dowchu Drukpa, Civil, DTU

53. Meta Analysis of Deep Learning Models for DoodleRecognition, ***8.Prateek Sethi***, ***8.Sachin Duhan***, ***8.Sanyam Gupta*** and ***3.Goonjan Jain***, Applied Mathematics, DTU

54. METAHEURISTIC BASED TECHNIQUES FOR OPTIMAL SIZING AND SITING OF DISTRIBUTEDGENERATIONS IN RADIAL DISTRIBUTION NETWORK, ***3.Prem Prakash***, Electrical, DTU

55. Models for Efficient Utilization of Resourcesfor Upgrading Android Mobile Technology, ***6.1.Abha Jain*** and ***6.1.Ankita Bansal***, CSE, DTU

56. Modified CCF based Shunt Active Power FilterOperation with Dead-Band Elimination for EffectiveHarmonic and Unbalance Compensation in 3-phase3-wire System, Alok Kumar Dubey, Jyoti Prakash Mishra and ***6.1.Amritesh Kumar***, Electrical, DTU

57. Modulation of polymer-basedimmobilized enzymes for industrialscale applications, ***7.Alka Dwevedi*** and Yogesh K. Sharma, Biotechnology, DTU

58. Operation and Control of Three-Phase Solar PVSystem Connected to Distribution Network, ***8.Hanish Mehla***, ***8.Manish Kumar***, ***8.Manish Yadav*** and ***3.Madhusudan Singh***, Electrical, DTU

59. Optimal Allocation of Capacitors in RadialDistribution Networks using Shannon's Entropy, ***8.Shubham Gupta***, ***3.Vinod Kumar Yadav*** and ***3.Madhusudan Singh***, Electrical, DTU

60. Parametric study of solar power tower driven combined supercritical CO2cycle and organic Rankine cycle using low GWP fluids, ***3.R.S. Mishra*** and ***6.Yunis Khan***, Mechanical, DTU

61. Performance Analysis of Optical CommunicationSystem for Various Filtering Operations UsingMIMO FSO System under Fog Condition, ***3.Anurag Chauhan***, ***8.Archana Kumari***, ***8.Priyanka Meena*** and ***8.Sanjay Kumar Meena***, Electronics, DTU

62. Performance Estimation of Power Factor CorrectedConverters for BLDC Motor Drive, ***8.Utkarsh Bhardwaj***, ***8.Abhishek Verma***, ***8.Sanjay Kumar*** and ***3.Saurabh Mishra***, Electrical, DTU

63. Potent Real-Time Recommendations UsingMulti-Model Contextual Reinforcement Learning, ***8.1.Anubha Kabra***, ***8.1.Anu Agarwal***, and ***3.Anil Singh Parihar***, CSE and Electronics, DTU

64. Quantitative EEG Feature Selection by MajorityVoting for Alcohol Use Disorder Detection, ***6.Ruchi Holker*** and ***3.Seba Susan***, IT, DTU

65. Recent Technological Advancement in Surrounding Gate MOSFETfor Biosensing Applications - a Synoptic Study, Amit Das, ***3.Sonam Rewari***, Binod Kumar Kanaujia and R. S. Gupta, Electronics, DTU

66. RPNDroid: Android Malware Detection usingRanked Permissions and Network Traffic, ***8.Madan Upadhayay***, ***8.Ashutosh Sharma***, ***8.1.Gourav Garg*** and ***3.Anshul Arora***, Applied Mathematics, DTU

67. Seasonal variability and kinetics of phosphate removalin a Phragmites-based engineered wetland, ***6.Harsh Pipi***, ***3.A. K. Haritash*** and Krishna R. Reddy, Environmental, DTU

68. Severity Classification of Diabetic Retinopathyusing ShuffleNet, ***8.Rohit Gambhir***, ***8.Shashank Bhardwaj***, ***8.Adwait Kumar*** and ***3.Ritu Agarwal***, IT, DTU

69. Solar Air Process Heating Systems for Drying of Agricultural Products, ***8.Rohit Kumar***, ***8.Suraj Kumar***, ***8.Sanchit Sharma*** and ***3.Dr. J.P. Kesari***, Mechanical, DTU

70. Structural, Electronic and NLO Properties of 6-aminoquinoline: A DFT/TD-DFT Study, Nupur Pandey, ***3.Mohan Singh Mehata***, Sanjay Pant and Neeraj Tewari, Applied Physics, DTU

71. Supply Chain Management Analysis in Online Retailers and Logistics, ***8.Siddhant Changkakati***, ***8.Shubhank Singh Chauhan***, ***8.Shivam Verma*** and ***3.Dr. N Yuvraj***, Mechanical, DTU

72. SUSTAINABLE MANUFACTURING – THE PAST, PRESENT AND FUTURE, ***8.Aryaman Sharma***, CSE, DTU

73. Synthesis, structural characterization and binding ability of A2B cobalt(III) corroles with pyridine***, 6.Jyoti***, Natalia Fridman, ***3.Anil Kumar*** and ***3.Sudhir G. Warkar***, Chemistry, DTU

74. The generalized second law of thermodynamics in viscous Ricci dark energy model, ***6.Ajay Kumar*** and ***3.C. P. Singh***, Mathematics, DTU

75. The link between employee career orientation, career management practices, citizenship behavior and gender: a moderated mediation model, ***3.Meha Joshi***, ***3.Girish Chandra Maheshwari*** and ***3.Rajan Yadav***, DSM, DTU

76. Thermodynamic and Comparative Analysis of Ejector Refrigeration Cycle and Absorption Refrigeration Cycle Integrated Wet Ethanol-Fueled HCCI Engine for Cogeneration of Power and Cooling, ***6.Mohd Asjad Siddiqui***, Abdul Khaliq and ***3.Rajesh Kumar***, Mechanical, DTU

77. Time-Borrowing Flip-Flop Architecture for MultiStage Timing Error Resilience in DVFS Processors, ***8.Avisekh Ghosh***, ***8.Mohd Saif Naseem*** and ***3.Chaudhry Indra Kumar***, Electrical, DTU

78. Time-Domain Characterization and Detection of Open Circuit Faults for H-Bridge Power Cell, ***3.Mayank Kumar***, Electrical, DTU

79. Tribological characterization of pultruded hybrid glass-jute fibre reinforced plastic composites from room temperature to 75°C, Navin Kumar, ***3.1.R.S. Walia*** and Surjit Angra, Production, DTU

80. Turing Machine based Syllable Splitter, ***8.Abhishek Mallik*** and ***8.Anavi Khetarpal***, CSE, DTU

81. Using the Boeing Max Air Disaster as A Role-play Scenario for Teaching Ethical Thinking, Ashish Hingle, ***8.1.Dr. Aditya Johri*** and Huzefa Rangwala, Mechanical, DTU

82. Utilizing Topic Modelling To Identify Abusive Comments On YouTube, ***8.Shubhanshu Shekhar***, ***8.Akanksha*** and ***8.Aman Saini***, CSE, DTU

83. Very Short-Term Photovoltaic (PV) Power Forecasting Using Deep Learning (LSTMs), ***8.Razzak Ahmad***, ***8.Pratyush*** and ***3.Rajesh Kumar***, Electrical, DTU

84. Wireless Energy Harvesting Cognitive Radio Network under Finite-Capacity Battery for Nakagami-m Fading, ***8.Nayanika Biswas*** and ***3.Piyush Tewari***, Electronics, DTU

85. wiseR: An end-to-end structure learning and deployment framework for causal graphical models, Shubham Maheshwari, ***8.1.Khushbu Pahwa*** and Tavpritesh Sethi, DTU

86. Wrong Side Vehicle Detection, ***Anirudh Goyal***, ***Antriksh Mangal*** and ***A.R.Kulkarni***, Electrical, DTU

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

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

2021 Fourth International Conference on Computational Intelligence and Communication Technologies (CCICT) | 978-1-6654-2392-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/CCICT53244.2021.00012

A Comprehensive Review of Adversarial Learning and Impact of Unsharing Weights Across Classes

S Vignesh
Delhi Technological University
Computer Engineering
New Delhi, Bawana Road
Email: s.vignesh_2k17co284@dtu.ac.in

Bhawna
Delhi Technological University
Computer Engineering
New Delhi, Bawana Road
Email: bhawna11ias@gmail.com

Shubham Anand
Delhi Technological University
Computer Engineering
New Delhi, Bawana Road
Email: s0.alively@gmail.com

Dr. Shailender Kumar
Delhi Technological University
Associate Professor
New Delhi, Bawana Road
Email: shailenderkumar@dce.ac.in

***Abstract*—It is evident from recent developments that deep learning has the potential to be a fundamental part in almost every new technology that comes in the future. Deep learning models have performed exceedingly well on standard image classification problems, but their performance drops drastically when presented with adversarial inputs that are created by adding specific small perturbations to the original image. This paper will be divided into two sections. The first section will be a complete review of the existing research in this field. We will provide the reader with the basic concepts of adversarial learning and a broad classification of various adversarial attacks and defenses. In the second section, we propose an ensemble model with max voting and test the impact of adversarial attacks by converting the 10-class problem over MNIST images into 10 binary classification problems. Each weight in the middle layers of a multi-class neural network is shared across all the output classes of the model. We call them the *shared weights* of the network. Our proposed model consists of no shared weights, shows a slight improve in accuracy against adversarial samples and can detect out-of-domain inputs.**

***Keywords*— adversarial learning, review, ensemble, shared weights**

I. Introduction

Deep learning has made quite a significant progress in the past decade in a wide domain of applications including healthcare, self-driving cars, robotics, etc. We already see neural networks being used in commercial self-driving cars, while algorithms like AlphaZero and AlphaGo defeat top human players of chess and Go respectively with ease. Since deep learning models seem ubiquitous, it is of utmost importance that we ensure their robustness since any compromise in their security can have a catastrophic impact on society.

Adversarial learning has been a topic of interest for researchers for over a decade now. It is seen that when slight specific perturbations are added to the input, the accuracy of neural networks drop drastically even though the noise added is so small that it is typically invisible to the naked eye. Over the years, it has become a game of seesaw where a defensive algorithm is proposed to counter an adversarial attack, but is broken by a new set of adversarial attacks. It is interesting to know that till date no defensive algorithm has been created which makes neural network robust against all kinds of adversarial attacks. In fact, most times a particular defensive method performs well only against the targeted adversarial attack it was made for.

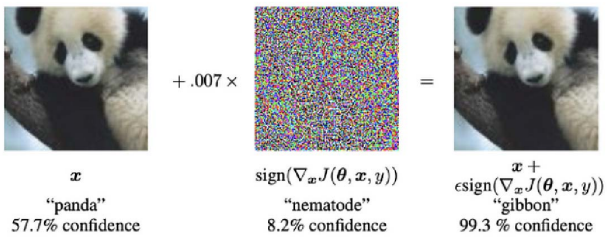


Fig. 1: A neural network network misclassifying a panda as a gibbon with high confidence after adversarial attack. Image credit to [Goodfellow et al. (2014)]

Before diving into discussions on adversarial attacks and defenses, let us clarify a few significant questions on adversarial attacks such as: What is adversary machine learning and what are adversarial inputs? How adversarial inputs are able to fool deep neural networks and what is meant by *robustness* of a neural network?

a) What is adversarial learning?:
Adversarial learning is the technique where a neural network is fooled by a deceptive input. The deceptive input differs very slightly from the original input and yet it was seen that the

978-1-6654-2392-2/21/$31.00 ©2021 IEEE
DOI 10.1109/CCICT53244.2021.00012

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:04:20 UTC from IEEE Xplore. Restrictions apply.

probability with which the model misclassified is quite high. The current struggle is to construct a defensive technique which is robust against all adversarial attacks.

b) What are adversarial inputs?:

An adversarial example is generated by adding a small (non-random) noise in the input data in order to fool the target neural network. A classic example of how an adversarial example looks like can be seen in Figure 1. It can be seen that after adding a perturbation with $\epsilon = 0.007$, the model incorrectly classified a "panda" as a "gibbon" with a lot of certainty.

Upon observation, it can also be noted that the perturbation added in the above example uses the sign of the gradient of the underlying model with respect to the input. The perturbation added to the input image is added in the direction of greatest loss to maximize the probability of misclassification. It is important to note that adversarial attacks are possible even when when the attacker does not have access to the gradients of the model. These are known as *black-box* attacks unlike *white-box* attacks when full access to the model is available.

c) How adversarial inputs are able to fool deep neural networks?:

It is interesting to wonder how even a small change in the input is capable of causing a change in prediction of the neural network. A deep neural network is highly non-linear and when an slight adversarial perturbation is added to the input image, it causes a large difference in feature learning in the intermediate hidden representations of the model and pushes the input image in the direction of the decision boundary of the true class and targeted class, making the model to misclassify the input.

d) What does it mean by adversarial robustness?:

In layman terms, adversarial robustness is a measure of how well a neural network is able to defend against an adversarial attack. Robustness is a relative term since a model which is robust against a particular adversarial attack might not hold against some other attack.

In the next section, we will discuss in detail some of the most successful adversarial attacks and defenses put forward by researchers over the years.

II. Adversarial attacks and defenses

In this section, we will classify different adversarial attacks, discuss some of the most successful adversarial attacks, and their impact on the accuracy of the model. Next, we will classify the different defenses proposed against these attacks and show improvement in adversarial performance. Throughout this section, we will use our base model for all evaluations which is defined below.

Base Model : A convolutional neural network with 2 convolution layers and 2 linear layers with dropout. Output space has 10 output classes trained with MNIST dataset for 30 epochs. This model has a test accuracy of 98.5% on the clean MNIST dataset, i.e., with no adversarial examples.

A. Classification of Adversarial Attacks

Broadly, adversarial attacks can be classified into white-box and black-box attacks as depicted in Table I. Black-box attacks can be further divided into three types depending upon the kind of environment of the target model. We go through each one of these classifications in detail.

1) White-box attacks: In this type of attack, the attacker has access to the parameters of the model. It is relatively easy to generate an adversarial example when the attacker has access to the gradient of the target model since the attacker can know the direction of maximum loss for the given input. Fast Gradient Sign Method (FGSM) (Goodfellow et al., 2014), Iterative Fast Gradient Sign Method (IFGSM) (Kurakin et al., 2016), Projected Gradient Descent (PGD) (Madry et al., 2017) are some of the most successful white-box attacks. White box attacks can be defended using *gradient masking*, i.e., by hiding the gradient of the model and making the model non-differentiable. White box attacks may become impractical in real-life scenarios when the model parameters are unavailable.

2) Black-box Attacks: The attacker does not have access to the model parameters in this type of attack. However, the attacker could make a copy of the targeted model and generate adversarial inputs via the cloned model. One of the reasons why adversarial examples are so hard to deal with is because of their ease of transfer, i.e., a single adversarial input may fool multiple models. In this case, the adversarial inputs generated for the cloned model are able to attack the targeted model whose gradients are actually unknown. Black-box attacks are widely classified into three major types as discussed below:

a) Transfer-based Attacks:

As discussed before, adversarial samples have the ability to transfer across models and hence attacks which make adversarial inputs to transfer to other models too are known as Transfer-based adversarial attacks. MIFGSM (Dong et al., 2018) makes a very good example for a transfer-based attack. Although it should be noted that MIFGSM is essentially a white-box attack since it needs the gradient of the model to generate adversarial samples but it need not be the parameters of the target model. An attacker can simply create a clone of the target model and generate adversarial examples using MIFGSM which will most likely fool the target model as well due to high transferability.

b) Score-based Attacks:

In these type of attacks, the attacker still does not have the model's parameters but they have access to the model itself, i.e., the adversary can hit queries to the model and get the class probability vectors or the labels as the output using which the attacker gets some information on the model parameters. In *Query-Based Setting* of score-based attacks, the attacker gets the list of probabilities of each class to which the input might belong. In *partial information Setting*, the attacker gets only the probabilities of top n classes which are most-likely to be the class in which the input belongs. In both these attack types, the attacker tries to estimate the gradient of the model using the probability predictions returned by the classifier for each class. Some examples of score-based black box attacks are ZOO (Zeroth order optimization)(Chen et al., 2017) and NES (Natural Evolutionary Strategies)(Zhou et al., 2020).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:04:20 UTC from IEEE Xplore. Restrictions apply.

TABLE I: Depicts a broad classification of adversarial attacks, their sub types and examples

Type of attack		Attack methods
White-box attacks		Eg. - FGSM (Goodfellow et al., 2014) - IFGSM (Kurakin et al., 2016) - PGD (Madry et al., 2017) - C&W (Carlini and Wagner, 2017b) - EOT (Athalye et al., 2018)
Black-box attacks	**Transfer-Based attacks**	Eg. - MIFGSM (Dong et al., 2018)
	Score-based attacks	Types: - Query limited setting (Ilyas et al. (2018),Suya et al. (2017)) - Partial information setting (Ilyas et al. (2017),Roponen et al. (2020)) Eg. ZOO (Chen et al. (2017),Chen and Gu (2020)) , NES (Wierstra et al., 2008)
	Decision-based attacks	Types: - Boundary attacks (Brendel et al., 2017) - Evolutionary attacks (Meunier et al. (2019),Zhou et al. (2020))

c) Decision-based attacks:
Attacks of this type only depend on the final output of the model such as label of the predicted class. Intuitively, attacking a model in a decision-based setting is relatively harder since there is little to no information about the parameters of the target model. In *Boundary-based attacks* (Brendel et al., 2017), *rejection-sampling* approach or also known as the *accept-reject* algorithm, is used. In this approach we start from a random adversarial input and traverse on the boundary separating the adversarial and non-adversarial input regions till the target image is achieved (Brendel et al., 2017). In *Evolutionary* attacks, attacks are made using evolutionary computation (Meunier et al. (2019),Zhou et al. (2020)).

B. Adversarial Attacks

In this section we will be discussing the technical details of some of the most successful adversarial attacks.

1) Fast Gradient Sign method (FGSM): FGSM was first introduced in (Goodfellow et al., 2014). FGSM is the simplest of adversarial attacks in which a slight perturbation is added to each pixel of the input image in the direction of the gradient of the loss function. The perturbation added is scaled by a factor of ϵ. Mathematically FGSM can represented as given in Equation (1):

$$X^{\text{adv}} = X + \epsilon \cdot \text{sign}(\delta_x J(X, Y)) \tag{1}$$

where ϵ is a very small value, X^{adv} is the generated adversarial image, X is the clean input image and δ_x is the gradient of the underlying neural network with respect to the original input X. it is seen that when the same adversarial input is given to different models, all of them misclassify but they agree on the class to which it belongs (Goodfellow et al., 2014).

Even though a neural network with no defenses against adversarial attacks can be fooled by adversarial inputs generated even by the simplest of attack algorithms, some adversarial attacks are stronger than others and are harder to defend.

It can be seen that the accuracy of the model drops drastically when no defenses are put up to hold adversarial attacks.

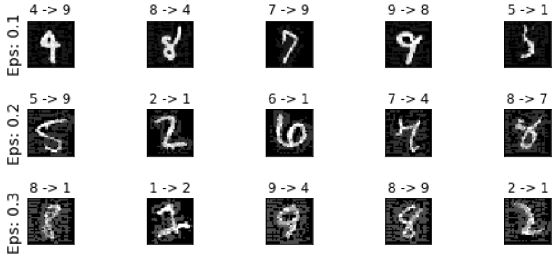


Fig. 2: Misclassified adversarial Images in FGSM attack from MNIST dataset for different values of ϵ

2) Iterative Fast Gradient Sign Method (IFGSM) and variants: IFGSM was introduced as an extension of FGSM in (Kurakin et al., 2016). It is also known by the name *Basic Iterative method*. This attack basically applies FGSM repeatedly for a certain number of iterations which is decided heuristically. The main idea is to increase the loss of the current class but in the same time not exceeding the ϵ L_∞ ball. IFGSM can be mathematically represented as given in Equation (2):

$$X_0^{\text{adv}} = X^{\text{adv}}, X_{N+1}^{\text{adv}} = \text{Clip}_{X,\epsilon}[X_N^{\text{adv}} + \epsilon \times \text{sign}(\delta_x J(X_N^{\text{adv}}, y)] \tag{2}$$

where *Clip* ensures that the perturbation in the image lies inside the ϵ L_∞ ball, X_{N+1} represents the adversarial input generated in the $(N+1)^{th}$ iteration. It should be noted that both FGSM and IFGSM do not make the classifier's prediction edge towards a specific class. They only increase the loss of predicting the correct class. For making the neural network to output a particular class of our choice, we use Targeted Fast Gradient Sign Method (TFGSM)(Xu et al., 2019) or *Iterative Least-Likely Class Method* (LL-class method)(Kurakin et al., 2016). The minor difference between the two is that LL-class method generates an adversarial input to favour prediction of the least-likely class whereas TFGSM can be used to favour the prediction in the direction of any class of the attacker's choice.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:04:20 UTC from IEEE Xplore. Restrictions apply.

As explained in (Kurakin et al., 2016), the least likely class is decided by using the prediction of the trained model. Intuitively, the class with the lowest probability in the probability vector of the model's prediction is chosen as the least-likely class. Mathematically, Iterative Least Likely Class Method (LL-class method), can be represented as given in Equation (3):

$$X_0^{\text{adv}} = X^{\text{adv}}, X_{N+1}^{\text{adv}} = \text{Clip}_{X,\epsilon}[X_N^{\text{adv}} + \epsilon \times \text{sign}(\delta_x J(X_N^{\text{adv}}, y_{LL})] \quad (3)$$

where y_{LL} is the least likely class. The change of operator from (+) to (-) while comparing with Equation (2) arises because the perturbation added to the input image should be in the direction of $\text{sign}(\log_p(y_{LL}|X))$ which is equivalent to $\text{sign}(-\delta_x J(X|y_{LL}))$ for models using cross-entropy loss function. Kurakin et al. (2016) also discussed in detail how input image transformations affect the impact of FGSM and its variants. After performing transformations such as changing image contrast and brightness, Gaussian blurring and JPEG encoding, they concluded that FGSM is the most robust to input image transformations along with the fact that changing brightness and contrast in the input images have negligible impact on the adversarial images.

Extending this further, (Dong et al., 2018) came up with a novel idea to enhance adversarial attacks using momentum. MIFGSM is very similar to FGSM apart from the fact that MIFGSM is more transferable than FGSM. When adversarial inputs were generated using IFGSM, where in every iteration FGSM was applied and pushed the adversarial input in the direction of greatest loss, the model started to overfit due to the presence of local maxima. Due to this overfitting, adversarial samples generated lost tranferability and were not successful in fooling black-box models (model parameters are not available to the attacker). It can be seen that our base model shows a similar drop in accuracy when exposed to both the attacks which was expected since for both attacks adversarial inputs were generated using the gradients of the target model itself and MIFGSM was primarily introduced for black box attacks.

3) Projected Gradient Descent (PGD): PGD is considered as a universal L_∞ attack (Madry et al., 2017). PGD is quite similar to IFGSM with the difference being that PGD initializes the adversarial input at a random point in the L_p ball around the input, does randomized initialization in every iteration and when $||x' - x||$ exceeds the L_p ball, it projects the perturbation back to the boundary of the L_p ball of the input. Figure 3[1] depicts the working of PGD using L_2-norm. It can be seen that the adversarial image is initialized randomly and then pushed towards the direction of highest loss in further iterations. The only drawback with PGD adversarial attack is that it requires more computational power since it calculates gradient at every iteration.

4) Carlini and Wagner Attack (C&W): C&W is also one of the most popular adversarial attacks existing due to the fact that it broke through defensive distillation which was the

[1]https://towardsdatascience.com/know-your-enemy-7f7c5038bdf3

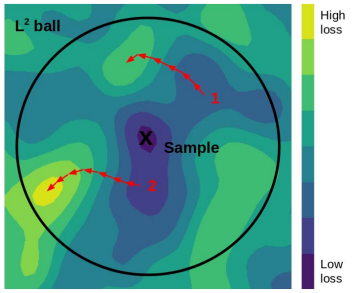


Fig. 3: PGD attack with random initialization of adversarial input in a region of low loss. In next iteration, PGD finds an adversarial input with high loss existing inside the L_2 ball

one of the strongest defenses against adversarial inputs at the time. (Carlini and Wagner, 2017b) converted adversarial input generation to an optimization problem as given below:

$$\begin{aligned} &\min D(x, x+\alpha) \\ &\text{s.t.} \quad C(x+\alpha) = t \\ &x + \delta \in [0,1]^n \end{aligned}$$

where x is the input image, D is the distance metric used (usually L_2-norm), C is the classifier and t is the target class. Solving this optimization problem will give the value of δ for which the classifier C will misclassify the input as the targeted class t. The problem arises because it is hard to solve this optimization problem with the given constraints. Hence, (Carlini and Wagner, 2017b) decided to modify the loss function by adding c times another loss function f. The constant c is chosen using binary search.Carlini and Wagner (2017b) gave a few choices for f, from which Equation (4) is widely used.

$$f(x') = \max(\max\{Z(x')_i : i \neq t\} - Z(x')_t, -k) \quad (4)$$

where $\max\{Z(x')_i\}$ is the log probability of the most-likely class and $Z(x')_t$ is the log probability of the target class. This ensures that the target class is slightly more likely than the current most-likely class. Carlini and Wagner attack is really difficult to defend against. In fact, (Carlini and Wagner, 2017a) tried to break 10 different defensive methods from 7 different published papers using C&W's *Zero-knowledge* attack. C&W attack can be used both as a white-box and black-box adversarial attack.

5) Expectation Over Transformation (EOT): (Athalye et al., 2018) showed that a image of a turtle when given to a image recognition model is misclassified as a "rifle" irrespective of the angle with which the picture was taken. It was seen that adversarial samples lost their effectiveness when the classifying model performs transformations to the input image. It basically uses alternating gradient steps and randomly generated transformations to account for naturally occurring image transformations. $D(x - x')$ is calculated only in the transformed points where D is the distance metric.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:04:20 UTC from IEEE Xplore. Restrictions apply.

TABLE II: Depicts Adversarial attacks and defenses and how they were countered

Attack/Defense	Countered by Defense/Attack	Brief description
Gradient Masking	Black-box Attacks (MIFGSM attack)	Using Transferability of adversarial inputs
Input Transformations Randomization	EOT (Expectation over Transformation)	Used the expectation value of the randomly generated gradients to generate adversarial inputs
FGSM attack IFGSM attack MIFGSM attack	Defensive Distillation	Used soft labels from the output of first network to train a second network. Probability distribution of each class gives more information than hard-labels
Defensive Distillation	Carlini and Wagner Attack	Modifying the optimization problem of adversarial input generation by using another objective loss function

C. Defenses against adversarial attacks

In this section, we will discuss the major classifications of defenses against adversarial attacks and some specific certified defenses which are theoretically proven to be robust against adversarial attacks given certain constraints.

We saw in the previous section, that all white box attacks used the gradient of the model and many black box attacks aimed to estimate the gradient of the model. Hence, the basic underlying agenda of a defensive technique against adversarial inputs will be to, by any means, hide the model's parameters from the attacker.

1) Adversarial training: Adversarial training is more of a brute force solution to our problem. The idea is simply to make the model learn to classify adversarial samples or detect them somehow by adding adversarial inputs in the training set itself. At first glance, this defense sounds quite robust but it is to be understood that adversarial inputs generated by each attack is different from each other and incorporating adversarial samples generated from all existing adversarial attacks becomes a tedious task. It is evident that this defense would require a lot of computational power. Instead, adversarial training is done to hold against a particular type of attack. Mathematically adversarial training can be written as given in Equation (5):

$$\min \max_{D(x,x')<\eta} J(\theta, x', y) \tag{5}$$

where D is the distance metric used, θ represents the network weights and $J(\theta, x', y)$ represents the adversarial loss function. The inner *max* expression works on making the best adversarial inputs and the outer *min* is to minimise the training loss.

We will be focussing on FGSM, IFGSM and MIFGSM training in this section. We will also be comparing the performance of all the mentioned attacks against FGSM, IFGSM and MIFGSM training.

FGSM Training Adversarial samples are generated using Equation (1) from images from the training data and the model is made to learn from them. Figure 4a shows the performance of FGSM, IFGSM and MIFGSM adversarial attacks after performing FGSM adversarial training.

MIFGSM Training Similarly, we train our base model to MIFGSM adversarial inputs and compare it's performance with other attacks. Figure 4c depicts the same.

From Figures 4a to 4c, we can conclude that IFGSM adversarial attack induces the greatest drop in accuracy irrespective of the type of adversarial training. Also, it can be noted that FGSM adversarial training improves the overall accuracy of the model against all three attacks.

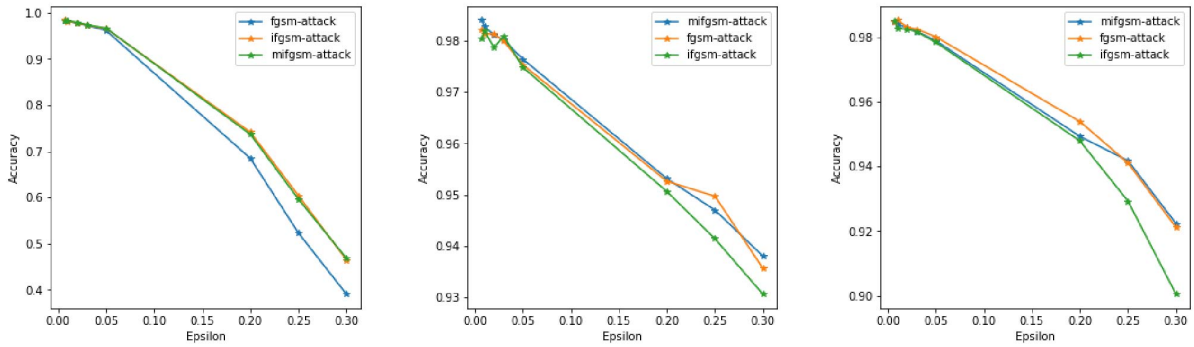
PGD Training PGD trained model have shown a significant increase in accuracy against many L_∞ attacks such as FGSM and PGD (He et al., 2016), but PGD training requires a lot of time and computational power. Also PGD adversarial training can only hold against L_∞ attacks and fails for other L_p-norm attacks such as C&W attack using L_2-norm (Carlini and Wagner, 2017b).

2) Input Transformations: In this defensive technique, a series of transformations is applied to the input to the model. Various input transformations which are effective in reducing the impact of adversarial inputs is discussed in (Guo et al., 2017). Image blurring, cropping and scaling and changing brightness of the image are some of the transformations used. It has been tested though that changing brightness and contrast in an input image doesn't affect the performance of the model (Guo et al., 2017). One drawback of these transformations is that it may cause *Gradient shattering*, i.e., gradients of the network become like white noise (Balduzzi et al., 2017), and hence making the model vulnerable to adaptive black-box attacks. This defense has been countered by EOT adversarial attack by simply estimating the gradient of the model (Athalye et al., 2018).

3) Randomization: A very important point to notice about neural networks is that they are mostly robust against random adversarial noise. It is only when adversarial images are generated specifically to increase the loss of the correct class when neural networks fail to perform accurately. Adding randomness to the neural network makes it hard for attackers to predict model's behaviour since random gradients are generated for every input. Adding a noise layer after each convolution layer (Liu et al., 2018) and adding random dropout layers in the network are some of the ways to add randomness to the model. Though any attack which can estimate the model's gradient using expectation value of random gradients generated can break this particular defense with ease.

4) Model Ensemble: Collection of various models integrated with different defensive techniques to hold against adversarial attacks together is known as an ensemble defense. Given an input to the model, each model produces an output of it's own

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:04:20 UTC from IEEE Xplore. Restrictions apply.



(a) With no adversarial training (b) With FGSM adversarial training (c) With MIFGSM adversarial training

Fig. 4: Measuring accuracy of our base model against adversarial attacks with and without adversarial training.

and the resultant output is calculated as an aggregation of all the outputs from the ensemble. Though it is impossible to make a strong ensemble defense against adversarial attacks using multiple weak defenses (He et al., 2017). Ensemble adversarial training (EAT) (Tramèr et al., 2017) models have shown high accuracy against black box adversarial attacks and performed much better than models with PGD training. (Pang et al., 2019) introduced a new training technique for ensemble models by promoting the *ensemble diversity*, i.e., diversity among the learned predictions of the ensemble members. It is called Adaptive Diversity Promoting (ADP)(Pang et al., 2019) training and has shown an increase in the accuracy of predictions of the ensemble.

5) Certified Defenses: Some defenses are experimentally proven to enhance robustness but do not have any theoretical guarantee of their effectiveness. Such defenses are called *Heuristic Defenses*. Whereas, defenses which are theoretically proven to be robust against their targeted attacks with the given constraints are known as *Certified Defenses*. In this section we will look at some of the certified defenses developed against adversarial attacks.

Input Gradient regularization

Input Gradient regularization is quite an old defensive technique and was introduced as a possible defense in Papernot et al. (2016) but is only currently looked in depth recently. If we consider deep learning models as a function, we see that for a given input x, we get an output y. When we see that in case of adversarial examples, a slight change in the input x ($x + \alpha$) causes a huge change in the output y. Hence, the concept of gradient regularization was re-introduced in (Ross and Doshi-Velez, 2018). This method is applicable to all differentiable models. Ross and Doshi-Velez (2018) argued that by making the input gradients smoother, the model will not only become more interpretable but also more robust to adversarial inputs and used *double back propagation* to smooth the input gradients. (Raghunathan et al., 2018) used a *Lipschitz* type approach to put an upper bound on the worst case loss of a neural network with 2 layers and proved that using their approach no L_∞ adversarial attack with an adversarial perturbation of $\epsilon = 0.1$ can make the model to misclassify more than 35% of the test samples on the MNIST dataset. Huster et al. (2018) lists the disadvantages of *Lipschitz smoothing* on complicated neural networks and argued against (Raghunathan et al., 2018) that representational drawbacks of Lipschitz smoothing will not allow complicated neural networks to be robust against adversarial attacks.

KNN based defenses K-nearest neighbor(KNN) algorithm works simply under the assumption that things which are similar to each other, stay together. For every new input X^{curr}, we calculate the distance $D(X^{\text{curr}} - X_i)$ where X_i is the ith sample from the dataset and D is the appropriate distance metric used. We sort the distances and pick a list of k labels which are closest to the current input and classify the input to the label which appears most times. In a similar way, (Papernot and McDaniel, 2018) introduced Deep k-Nearest Neighbors(DkNN) which used nearest neighbor algorithm for every layer of the neural network and made sure that for a given test input the intermediate calculation results in each layer are in sync with the final prediction of the model. If the intermediate features learnt by the model for a given test input are very different from the features learnt in the training phase with inputs sharing the same final prediction, then the prediction by the model is termed as *abnormal*. (Wang et al., 2018) discussed the robustness of KNN algorithm against adversarial attacks and proposed an efficient 1-nearest neighbor by simply noticing that 1-nearest neighbor algorithm fails only for outliers, i.e., when oppositely labelled data samples are close to each other. The proposed algorithm (Wang et al., 2018), eliminates the outliers and outperformed other defenses when exposed to attacks with large adversarial perturbations.

III. Ensemble model via weight unsharing

All the weights before the penultimate layer of the neural network are shared across all the output classes of the network. A weight is just a number and during the training phase of a neural network, each weight adapts itself to favour all classes in the output space. In this section, we explore in brief the benefits and drawbacks of shared weights in a network and

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:04:20 UTC from IEEE Xplore. Restrictions apply.

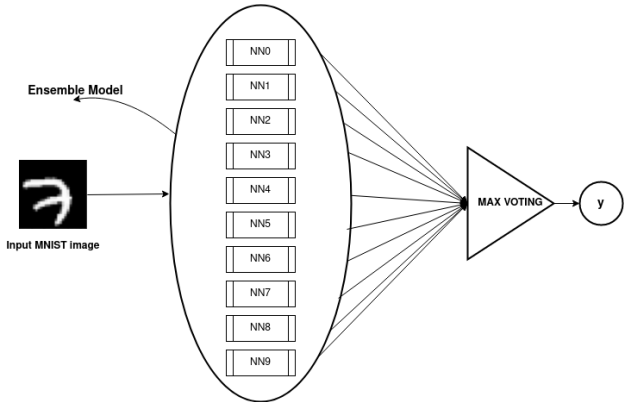


Fig. 5: Figurative representation of the proposed ensemble model

how they affect the performance of the model. To work on this experiment, we will be using the MNIST dataset and to minimize weight sharing, we propose the following model:

Ensemble Model: Instead of using a neural network with 10 output classes to predict MNIST digits, we will be using an ensemble model of 10 neural networks each having only 2 output classes. Each neural network consists of 2 convolution layers and 2 linear layers and is trained to recognize a particular digit from the MNIST dataset via the *one vs rest* approach. There is no weight sharing in this model since each member of the ensemble is trained to identify only one MNIST digit. This model achieves a test accuracy of 98.5% on a clean MNIST dataset after a training time of just 7 epochs. Use Figure 5 for reference. We will now explain the working of our model in detail:

1) Each neural network in the ensemble has two output classes. Training is done in such a way that the output from NN_i (Figure 5) will be a vector of size 2 in which the value in the first index will be a probability with which the given MNIST input is *not* i and the value in second index will indicate the probability with which the given input *is* i according to NN_i.
2) An input image to the ensemble model will pass through each of the 10 internal neural networks. The output from the ensemble will be 10 class probability vectors, each of size 2.
3) The final prediction will be calculated using *max voting*. The final probability (P) from the ensemble will be the maximum of 10 decimal values by taking the second index of each of the 10 class probability vectors obtained.
4) The final prediction of the ensemble will depend on the final probability (P). If $P >= 0.5$ and P is obtained from NN_k where $k \in [0, 9]$, then the final prediction will be k. Whereas if $P < 0.5$, the image is discarded as either an adversary or not an MNIST digit.

We now state the benefits of using such an ensemble model against adversarial attacks:

1) Our model will be able to detect out of domain inputs, i.e., images which are not MNIST digits, with ease unlike normal multi-class models which will misclassify any input image as one of the MNIST digits.
2) Since the model will discard the input image if the probability obtained after applying max voting falls below a certain threshold P_t ($P_t = 0.5$ in our proposed model), the attacker will have to ensure that the confidence with which the model misclassifies crosses that threshold. Also, instead of attacking a single model in case of multi-class models, the attacker will have to attack two models, i.e., increase the loss of the neural network corresponding to the true class of the input and decrease the loss of the neural network corresponding to the target class.
3) From Figure 6, it can be seen that our proposed ensemble model outperforms our base model when attacked with FGSM, IFGSM and MIFGSM adversarial samples generated using the base model.

Even though this model shows certain improvement against adversarial inputs, it is quite expensive computation-wise. Due to a large number of weights, training time is really high and the model requires a lot of computational power to train even for a few epochs.

IV. Conclusion

With this paper, we hope to give the reader a thorough understanding on the workings of adversarial samples and why they have the potential to be a great threat in the future. With a lot of research currently existing on the topic, we provide the reader a fairly detailed knowledge on the basic concepts in adversarial learning, a broad classification of adversarial attacks and defenses, a walk through history by mentioning the existing adversarial attacks and defenses and how they were countered and finally a brief study on shared weights in a neural network and how they impact the decision making of a model.

V. Future Scope

In the coming future, we will be exploring more deeply on the impacts of shared weights in complicated deep neural networks and adversarially trained models. We will also be shifting our attention to input gradient regularization and various smoothing techniques since they currently the meta idea to defend against adversarial attacks.

References

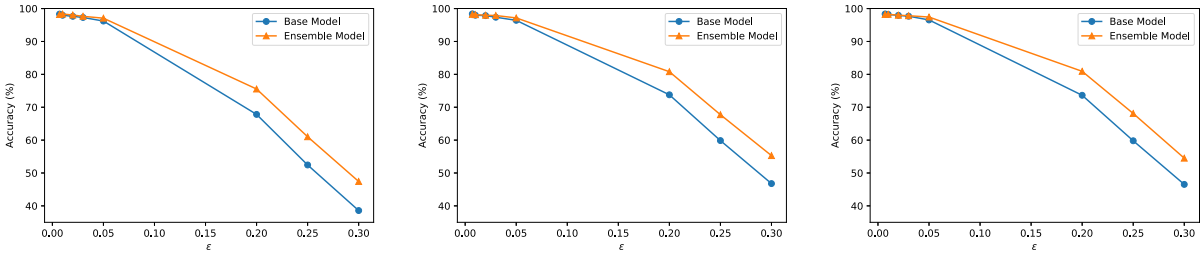
A. Athalye, L. Engstrom, A. Ilyas, and K. Kwok. Synthesizing robust adversarial examples. In *International conference on machine learning*, pages 284–293. PMLR, 2018.

D. Balduzzi, M. Frean, L. Leary, J. Lewis, K. W.-D. Ma, and B. McWilliams. The shattered gradients problem: If resnets are the answer, then what is the question? In *International Conference on Machine Learning*, pages 342–350. PMLR, 2017.

W. Brendel, J. Rauber, and M. Bethge. Decision-based adversarial attacks: Reliable attacks against black-box machine learning models. *arXiv preprint arXiv:1712.04248*, 2017.

N. Carlini and D. Wagner. Adversarial examples are not easily detected: Bypassing ten detection methods. In *Proceedings of the*

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:04:20 UTC from IEEE Xplore. Restrictions apply.



(a) against FGSM adversarial attack (b) against IFGSM adversarial attack (c) against MIFGSM adversarial attack

Fig. 6: Impact of FGSM, IFGSM and MIFGSM attacks on proposed ensemble model. The proposed model outperforms the base model by at least 9%

10th ACM Workshop on Artificial Intelligence and Security, pages 3–14, 2017a.

N. Carlini and D. Wagner. Towards evaluating the robustness of neural networks. In *2017 ieee symposium on security and privacy (sp)*, pages 39–57. IEEE, 2017b.

J. Chen and Q. Gu. Rays: A ray searching method for hard-label adversarial attack. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1739–1747, 2020.

P.-Y. Chen, H. Zhang, Y. Sharma, J. Yi, and C.-J. Hsieh. Zoo: Zeroth order optimization based black-box attacks to deep neural networks without training substitute models. In *Proceedings of the 10th ACM workshop on artificial intelligence and security*, pages 15–26, 2017.

Y. Dong, F. Liao, T. Pang, H. Su, J. Zhu, X. Hu, and J. Li. Boosting adversarial attacks with momentum. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 9185–9193, 2018.

I. J. Goodfellow, J. Shlens, and C. Szegedy. Explaining and harnessing adversarial examples. *arXiv preprint arXiv:1412.6572*, 2014.

C. Guo, M. Rana, M. Cisse, and L. Van Der Maaten. Countering adversarial images using input transformations. *arXiv preprint arXiv:1711.00117*, 2017.

K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

W. He, J. Wei, X. Chen, N. Carlini, and D. Song. Adversarial example defense: Ensembles of weak defenses are not strong. In *11th {USENIX} workshop on offensive technologies ({WOOT} 17)*, 2017.

T. Huster, C.-Y. J. Chiang, and R. Chadha. Limitations of the lipschitz constant as a defense against adversarial examples. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 16–29. Springer, 2018.

A. Ilyas, L. Engstrom, A. Athalye, and J. Lin. Query-efficient black-box adversarial examples (superceded). *arXiv preprint arXiv:1712.07113*, 2017.

A. Ilyas, L. Engstrom, A. Athalye, and J. Lin. Black-box adversarial attacks with limited queries and information. In *International Conference on Machine Learning*, pages 2137–2146. PMLR, 2018.

A. Kurakin, I. Goodfellow, S. Bengio, et al. Adversarial examples in the physical world, 2016.

X. Liu, M. Cheng, H. Zhang, and C.-J. Hsieh. Towards robust neural networks via random self-ensemble. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 369–385, 2018.

A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu. Towards deep learning models resistant to adversarial attacks. *arXiv preprint arXiv:1706.06083*, 2017.

L. Meunier, J. Atif, and O. Teytaud. Yet another but more efficient black-box adversarial attack: tiling and evolution strategies. *arXiv preprint arXiv:1910.02244*, 2019.

T. Pang, K. Xu, C. Du, N. Chen, and J. Zhu. Improving adversarial robustness via promoting ensemble diversity. In *International Conference on Machine Learning*, pages 4970–4979. PMLR, 2019.

N. Papernot and P. McDaniel. Deep k-nearest neighbors: Towards confident, interpretable and robust deep learning. *arXiv preprint arXiv:1803.04765*, 2018.

N. Papernot, P. McDaniel, X. Wu, S. Jha, and A. Swami. Distillation as a defense to adversarial perturbations against deep neural networks. In *2016 IEEE symposium on security and privacy (SP)*, pages 582–597. IEEE, 2016.

A. Raghunathan, J. Steinhardt, and P. Liang. Certified defenses against adversarial examples. *arXiv preprint arXiv:1801.09344*, 2018.

J. Roponen, D. R. Insua, and A. Salo. Adversarial risk analysis under partial information. *European Journal of Operational Research*, 287(1):306–316, 2020.

A. Ross and F. Doshi-Velez. Improving the adversarial robustness and interpretability of deep neural networks by regularizing their input gradients. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

F. Suya, Y. Tian, D. Evans, and P. Papotti. Query-limited black-box attacks to classifiers. *arXiv preprint arXiv:1712.08713*, 2017.

F. Tramèr, A. Kurakin, N. Papernot, I. Goodfellow, D. Boneh, and P. McDaniel. Ensemble adversarial training: Attacks and defenses. *arXiv preprint arXiv:1705.07204*, 2017.

Y. Wang, S. Jha, and K. Chaudhuri. Analyzing the robustness of nearest neighbors to adversarial examples. In *International Conference on Machine Learning*, pages 5133–5142. PMLR, 2018.

D. Wierstra, T. Schaul, J. Peters, and J. Schmidhuber. Natural evolution strategies. In *2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence)*, pages 3381–3387. IEEE, 2008.

J. Xu, Z. Cai, and W. Shen. Using fgsm targeted attack to improve the transferability of adversarial example. In *2019 IEEE 2nd International Conference on Electronics and Communication Engineering (ICECE)*, pages 20–25. IEEE, 2019.

Y. Zhou, Y.-a. Tan, Q. Zhang, X. Kuang, Y. Han, and J. Hu. An evolutionary-based black-box attack to deep neural network classifiers. *Mobile Networks and Applications*, pages 1–14, 2020.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:04:20 UTC from IEEE Xplore. Restrictions apply.

Multimedia Tools and Applications
https://doi.org/10.1007/s11042-021-11415-9

A deep multimodal network based on bottleneck layer features fusion for action recognition

Tej Singh[1] · **Dinesh Kumar Vishwakarma**[1]

Received: 13 September 2020 / Revised: 25 January 2021 / Accepted: 10 August 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Human Activity Recognition (HAR) in videos using convolution neural network become the preferred choice for researcher due to the tremendous success of deep learning models for visual recognition applications. After the invention of the low-cost depth sensor, multiple modalities based activity recognition systems were successfully developed in the past decade. Although it is always challenging to recognize the complex human activities in videos. In this work, we proposed a deep bottleneck multimodal feature fusion (D-BMFF) framework that fused three different modalities of RGB, RGB-D(depth) and 3D coordinates information for activity classification. It helps to better recognize and make full use of information available simultaneously from a depth sensor. During the training process RGB and depth, frames are fed at regular intervals for an activity video while 3D coordinates are first converted into single RGB skeleton motion history image (RGB-SklMHI). We have extracted the features from multimodal data inputs using the latest deep pre-trained network architecture. The multimodal feature obtained from bottleneck layers before the top layer is fused by using multiset discriminant correlation analysis (M-DCA), which allows for robust visual action modeling. Finally, using a linear multiclass support vector machine (SVM) method, the fused features are categorized. The proposed approach is evaluated over four standard RGB-D datasets: UT-Kinect, CAD-60, Florence 3D and SBU Interaction. Our framework produces outstanding results and outperformed the state-of-the-art methods.

Keywords Human Activity Recognition (HAR) · Deep learning · DCA · SVM

✉ Dinesh Kumar Vishwakarma
dvishwakarma@gmail.com

Tej Singh
ttomar07@gmail.com

[1] Biometric Research Laboratory, Department of Information Technology, Delhi Technological University, Delhi 110042, India

Published online: 20 August 2021

1 Introduction

Today, human activity recognition in videos has become an imperative choice for computer vision researchers due to its wide variety of real-life applications. It includes e-health, patients monitoring, assistive daily living activities, video surveillance, security and behavior analysis, sports analysis and many more [1]. One of the ultimate goals of the artificial intelligence society is to create an automatic system that can accurately recognize and interpret human actions. However, human activity recognition is still a difficult problem in machine learning and many primary problems remain unanswered, such as intra-class variation, varying illuminance, occlusion and similarity in acts. The primary objective recognition system is robust modeling of action and competent representation. Unlike an image, the modeling of human behavior in video sequences is based on the representation of spatio-temporal characteristics. The spatiotemporal models extracted the feature of spatial appearance present in video frames and existing variations of the pose. In this context, the spatial features in handcrafted approaches for activity recognition is relied on textural information while optical flow estimated the pose variations in the video sequences [43, 44]. After the tremendous progress of deep learning architectures on human pose estimation, object detection, segmentation, speech analysis, and object tracking [36, 37]. The deep learning model also plays a central role in visual recognition tasks. Unlike traditional solutions, deep learning-based approaches provide a new way to automatically extract the features from images.

With the invention of advanced Kinect depth sensor various methods based on RGB, depth and 3D coordinates and their various combinations are introduced in the literature [17, 18, 26, 34]. However, very few approaches are available based on the combination of these three modalities. These works have shown promising results but had drawbacks due to error present in joint locations body parts used to represent action [2, 9, 20]. In the proposed work, we utilized three modalities RGB, depth, and 3D coordinates information for activity classification because it helps for better recognition and complete utilization of information available from a depth sensor video simultaneously. It is observed that sometimes skeleton data is degraded due to noise and occlusions present in depth images and less effective for representation. Therefore, motivated by the work [27], we utilized the concept of color motion history based skeleton images(RGB-SklMHI). These skeleton images are obtained by averaging the binary MHIs on the sequential time interval for each action video sequence.

In recent years, various fusion methods have been applied to machine learning. In this context, widely adopted methods are early fusion, late fusion and decision level fusion [28, 34]. Due to robust discriminative features extraction, the feature level fusion is more dominated among these techniques. It is noted that different characteristics can provide body descriptions with various aspects and constitute additional information. We cannot only preserve the efficient multi-feature information by integrating these characteristics but also eliminate the redundant information to some extent. The typical features fusion techniques are canonical correlation analysis (CCA) [38], and discriminant correlation analysis (DCA) [12]. We used a multiset DCA technique to combine the multimodal features derived from deep pre-trained architecture in the proposed work. Further, the multimodal fused features are categorized using multiclass SVM classifier [6]. The main contributions to the proposed work are outlined in a nutshell as follows.

- We implemented a hybrid multimodal feature deep CNNs system (D-BMFF) that used RGB, depth, and 3D co-ordinate to incorporate full use of depth sensor information for activity recognition in videos.
- We have extracted bottleneck feature (intermediate layers) by fine-tuning the latest pre-trained deep architectures for action modelling to avoid the overfitting at top layers.
- The extracted deep feature is combined with the multiset DCA technique from three different data streams and the class of actions is classified by the SVM machine learning technique.
- On four distinct standard human activity datasets, we apply the proposed framework and obtain promising results. It demonstrated the effectiveness of the approach and the applicability to the tasks of video classification.

This work is formulated as follows: Sect. 1 consisted of introduction, applications and challenges in HAR. In Sect. 2, studied some related methods based on handcrafted features and deep learning for the recognition of human activity in videos. In Sect. 3, the proposed D-BMFF framework is explained in detail. For the evaluation of the HAR framework, Sect. 4 consists of the training setting and implementation details. The findings and comparison with similar works of human activity datasets have been discussed in Sect. 5. At last, the work has been concluded in Sect. 6.

2 Related work

In this section, we reviewed work related to the recognition of human activity in RGB-D datasets. Traditional handcrafted features solutions have shown promising results, but have relied more on action classification features descriptors. These solutions required more work and expertise in the field of knowledge. On the other hand, deep learning approaches are dominated by the automatic extraction of features from raw videos. Existing solutions based on the modality used for HAR can be divided into four parts: 1) RGB-based approach, 2) depth-based approach, 3) a skeleton-based approach, and 4) joint modality-based approach.

Various methods using RGB videos have been proposed in the literature over the last two decades. These solutions are based on global and local features descriptors to represent action classes such as motion history images (MHIs) [4, 5], spatiotemporal interest features (STIP) [22], optical flow [7], and improved dense motion trajectories [44, 45]. After the success of deep CNN model in image classification tasks, Ji et al. [16] suggested a deep architecture based on 3D CNN to distinguish actions using RGB frames. Simoyan and Zisserman [34] proposed a two-stream network using RGB frames and optical flow to recognize human actions in videos. Various solutions based on depth frames have been introduced to identify the action class in the video [3, 19].

In skeleton-based solutions, action representation is proposed using 3D skeleton coordinates. In this context, Huynh et al. [14] proposed a spatiotemporal approach using skeleton joints based on Pachinko Allocation Model for action recognition. Devanne et al. [8] proposed a spatio-temporal model using motion trajectories in Riemannian manifold space. Each joint of the skeleton was represented as a motion trajectory in 3D action space. But their approach is limited to similar and one-person activities. Vemulapalli et al. [42] have proposed a 3D space group model based on 3D relative geometry between the different body parts for skeleton action representation.

They modeled human actions in 3D lie groups, and classification is performed using dynamic time wrapping and vector support. However, their approach is limited to the classification of single-person activity only. Yang et al. [47] have introduced a latent max-margin multitasking learning approach to skeleton data recognition. Their skelets model has learned and correlates the mid-level granularity of the joints to represent action classes. Salih and Youssef [30] have introduced a spatiotemporal model based on spherical angles to represent the static pose and joint displacement of a given action skeleton sequence. Li et al. [23] addressed the key pose based human activity recognition using shape dynamic time wrapping. They utilized segmented skeleton feature sequence into a key pose and potential difference for an atomic motion for activity recognition. Liu et al. [25] proposed a HAR model based on the pose specificity metric to identify the activity classes in order to address the problem of broad variability among the action classes using skeleton data. A weighted histogram 3D joint skeleton approach was introduced by Shabaninia et al. [32] to resolve the identification of human action. In addition, to describe the temporal variance and action, a weighted motion energy function is used. Ghodsi et al. [11] proposed a spatio-temporal action template based on the temporal analysis of action samples for the recognition of human activity using 3-D skeleton data. Action classes are represented as multi-dimensional signals and are capable of dealing with current variations in activities such as speed.

Details on shape and motion can be found to be correlated with in-depth sequences, but both of these are hard to catch simultaneously. In order to compensate for the limitations of an individual modality, hybrid feature approach introduced by Nguyen et al. [26]. A Gaussian descriptor based on high-order statistics of local features at two levels used to represent actions. They used K-means and sparse coding technique to compensate for the information loss caused by heterogeneous vector features obtained from two separate input streams: depth and skeleton. Kong et al. [21] have introduced a multi-modal fusion-based feature learning approach to represent the local dynamics of each joint action class. The classification of the action shall be carried out using a max-margin framework. Raman and Maybank [29] presented a two-level hierarchical model based on the Hidden Markov Model for activity recognition with depth sequences and skeleton coordinates. They provided flexible representation in order to discriminate against action classes with similar activity labels. Khaire et al. [20] proposed a three stream deep model for recognized human activity. They built MHI from RGB, deep motion map sequences, and average skeleton images to train CNN's model and extracted features are merged at decision level for final action classification.

3 The proposed framework for action recognition

It is observed that when a new model is trained from raw data, a deep learning model requires a lot of training data and computing power. As a result, it is not a feasible task to collect and label a large amount of data that makes it difficult to apply a deep learning model. We used deep transfer learning for human activity recognition because of limited training data.

Three modalities including RGB, depth, and 3D skeleton coordinates for activity recognition have been used in this proposed work. We fine-tuned the bottleneck layer (Conv '7b') and the fully connected (FC) layer of the pre-trained Inception-Resnet-v2(IR-v2) architecture [41] to extract features from three separate input streams. At regular intervals of ten frames, the RGB and depth frames are fed to deep neural network for extraction of spatial features and a single Skel-MHI image for each activity is given as an input for evidence of additional temporal features. In addition, the multimodal feature extracted from three data streams is combined using the Multiset DCA fusion technique at (Conv'7b ') and (FC) layers. The activity classification is carried through the SVM machine learning technique. Using deep feature representation, multiset DCA, and SVM, the human activity recognition framework is shown in Fig. 1. In the proposed work, we have conducted the experiment on different state-of-the-art deep architectures. Further, classification results produced by the individual data streams are evaluated on standard datasets.

4 Deep features extraction

Deep inception architecture [40] has shown promising results with little computational cost compared to other deep architectures. It can be seen from work [41] that in recent ILSVRC challenges, the addition of a residual unit with the conventional deep model has provided a state-of-the-art efficiency. In addition, it has increased the training speed of the Inception model by a sufficient margin and significantly improved the task of image recognition. In the proposed framework, the RGB and depth frames are fed to the pre-trained network at regular intervals of ten frames for extraction of spatial features and a single Skel-MHI image is provided as an input for additional evidence of temporal features for each activity. Deep IR-v2 architecture is used for the extraction of

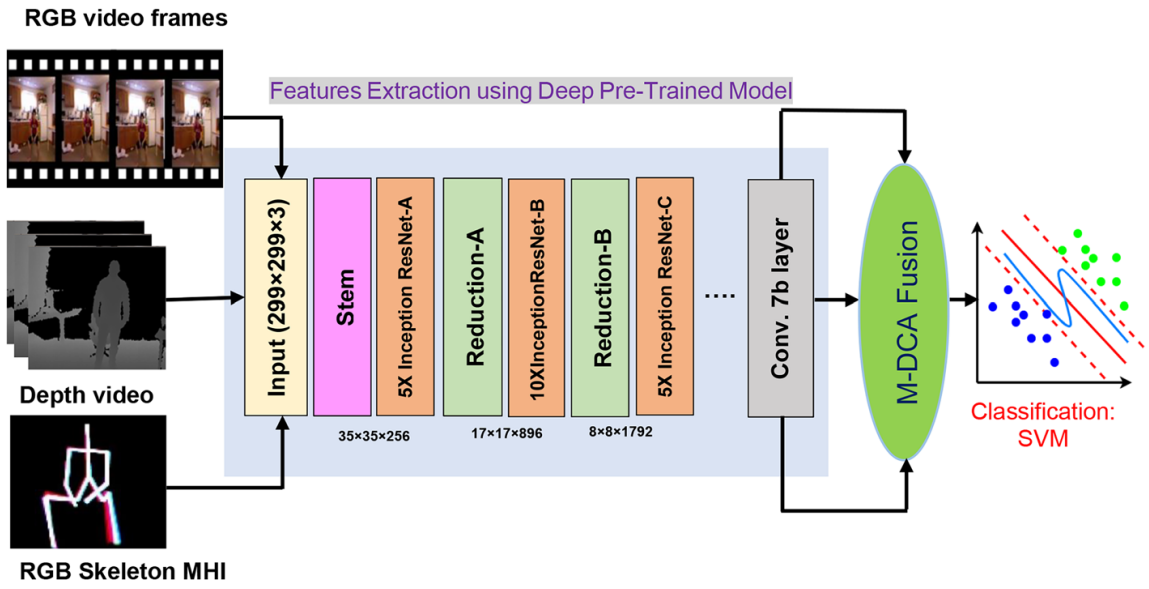


Fig. 1 The human activity recognition framework is shown using deep feature representation, multiset DCA, and SVM

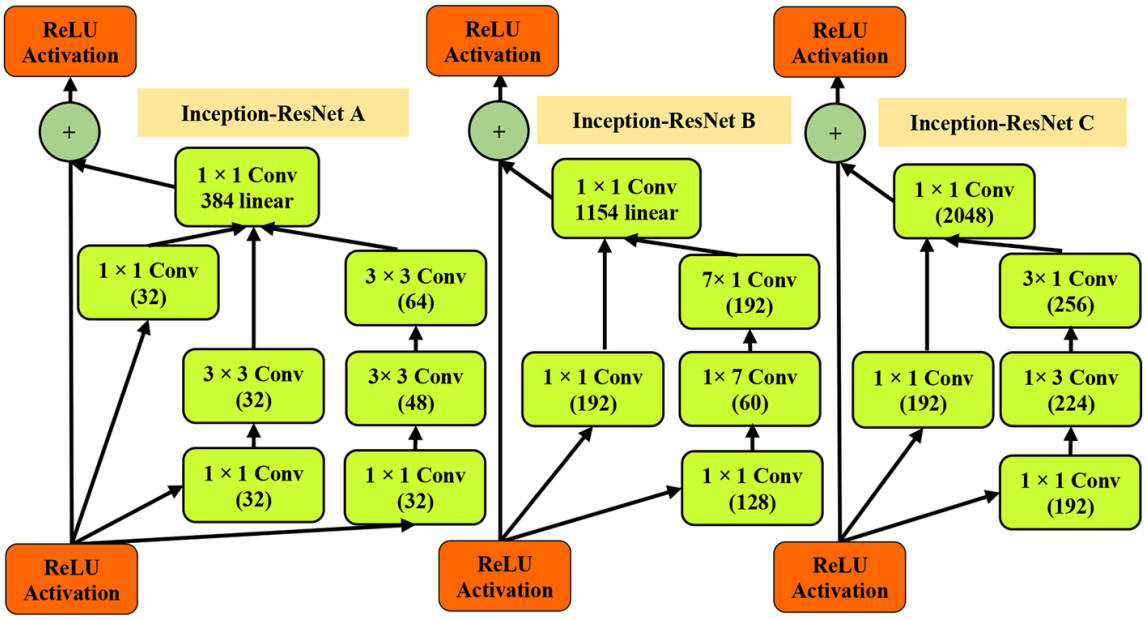


Fig. 2 Shows the systematic block diagram for 35x35 grid module of Inception ResNet-A, 17x17 grid module of Inception ResNet-B, and 8x8 grid module of Inception ResNet-C

features and fine-tuned its last few layers for the classification of activities. This deep model is trained in a large image database and has shown excellent results in the image classification with minimal error compared to other deep models. This network is 164 layers deep and accepts $299 \times 299 \times 3$ input images. It comprises two deep residual connections architectures [13] and the latest Inception-v4 [41] inception architecture. It is noted that residual connections are a preferable choice for training very deep architecture and will increase the computational efficiency of deep Inception by adding the residual units to the architecture. The detailed structure of blocks A, B and C of the Inception-ResNet is shown in Fig. 2. The three blocks of the inception block are followed by the expansion layer of the filter (1×1 convolution layer without activation). This 1×1 Conv layer helps to scale the size of the filter bank in addition to matching the depth of the input.

4.1 Skeleton motion history images (RGB-SklMHI)

The concept of creating an MHI color skeleton from 3D coordinates is introduced in the work Phyo et al. [27]. It is observed that 3D joint coordinates are not sufficient for the representation of action in video frames because of self-occlusion of upper body and noise. Furthermore, it is found that binary MHI images are not sufficient to discriminate against similar patterns of motion, such as 'sitting down' and 'standing up' position. Therefore, images of color skeleton MHIs are obtained from joint coordinates based on the average of binary MHIs for each action class in the sequential time interval. The RGB-SklMHI is based on a temporal view template and is robust in terms of noise, occlusion and illumination. The formation of RGB-SklMHI from 3D coordinates based on the sequential time interval is described in Algorithm 1. In the proposed framework,

we used SklMHI images for each dataset obtained with the procedure explained in Algorithm 1. These images are given as an input to a deep model to extract patterns of temporal variations in action videos. The deep feature from RGB-SklMHI is extracted using deep IR-v2 model and merged with other modalities concept for robust representation of the action.

Algorithm 1: Colorization of Skeleton Motion History Image

Input: 3D joints coordinates

Output: RGB Skeleton MHIs

Input:

1. *N=1,2,3……., n ; total joints*
2. *f=1,2,3, ……,n ; total frames*
3. *p=1,2,3, ...,n ;all labels*
4. *I (X, Y) template RGB-SklMHI*
5. *I(SkeletonConnectionMap)= [spine, left, right, …….]*
6. *I(ColorSet)= [0,1,0;0,1,0.33, ……..]*
7. *RGBSkeleton Img ( )*

Procedure: *Create_SklImg ((N, p, f, I (X, Y), ColorSet, color_c, SkeletonConnectionMap)*

for *p ←all labels* ***(do)***

for *f ←all frames* ***(do)***

for *frame= startframe +(N-1): endframe*
I(X,Y)=zeros(height,width,3);
color_c=1;
for *k = frame-(N-1): frame*
color_img = (ColorSet, color_c, SkeletonConnectionMap);
I(X,Y)= I(X,Y)+ color_img;
color_c= color_c + 1;
end for
end for
end for
end for
end procedure
Output: *RGB Skeleton MHIs.*

4.2 Feature fusion using discriminate correlation analysis (M-DCA)

Discriminant correlation analysis (DCA) feature fusion technique is proposed by Haghighat et al. [12] . DCA technique is useful for multimodal fusion, and real-time biometric recognition system. DCA is a compelling feature fusion technique that maximizes pairwise correlations between the features set. It also removes the correlation between classes and restricts the correlation to be within classes with low computational complexity. Feature vectors extracted from a deep model with a high dimension are observed. In this context, DCA is an effective feature fusion approach with a lower dimensional feature vector while combining the different feature vectors. The key steps for the fusion of feature sets using the DCA technique are discussed in Algorithm 2.

Algorithm 2	**Muti-set features fusion using DCA**
Step 1	The mean $\widetilde{m}_i$ and $\widetilde{m}$ is calculated for each class and training data, respectively as: $\widetilde{m}_i = \frac{1}{n_i}\sum_{j=1}^{n_j} m_j^i$, $\widetilde{m} = \frac{1}{n}\sum_{i=1}^{c} n_i \widetilde{m}_i$, where, n number of images, c number of classes.
Step 2	The covariance matrix is computed as $\mathrm{C} = \phi^{\mathrm{T}}\phi$ where $\phi = \left(\sqrt{n_i}(\widetilde{m}_i - \widetilde{m}), \ldots\ldots\ldots, \sqrt{n_c}(\widetilde{m}_c - \widetilde{m})\right)$,
Step 3	Calculate the SVD of C as $\mathrm{C} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{T}}$, $\boldsymbol{\Lambda} = diag(\lambda_1\lambda_{2,\ldots\ldots\ldots\ldots}\lambda_c)(\lambda_1 \geq \lambda_{2,\ldots} \geq \lambda_c)$, Where λ_i represent eigenvalue of C and i^{th} column of $\boldsymbol{U}$ is the corresponding eigenvector.
Step 4	The transform matrix is calculated as $\boldsymbol{P} = \boldsymbol{\phi}\boldsymbol{U}_r\boldsymbol{\Lambda}_{r\times r}^{\frac{-1}{2}}$.
Step 5	Repeat the steps (1-4) to compute the transform matrix for $\boldsymbol{\mathcal{X}_{RGB}}, \boldsymbol{\mathcal{X}_{RGB-D}}$, and $\boldsymbol{\mathcal{X}_{Skel-MHI}}$ streams separately.
Step 6	Calculate the transform data: $\boldsymbol{\mathcal{Z}_1} = \boldsymbol{P_1^{\mathrm{T}}}\boldsymbol{\mathcal{X}_{RGB}}$, $\boldsymbol{\mathcal{Z}_2} = \boldsymbol{P_1^{\mathrm{T}}}\boldsymbol{\mathcal{X}_{RGB-D}}$, and $\boldsymbol{\mathcal{Z}_3} = \boldsymbol{P_1^{\mathrm{T}}}\boldsymbol{\mathcal{X}_{Skel-MHI}}$.
Step 7	Evaluate the two-set covariance matrix of the transformed feature set as: $\boldsymbol{S_{bwn}} = \boldsymbol{\mathcal{Z}_1}\boldsymbol{\mathcal{Z}_2}^{\boldsymbol{T}}$.
Step 8	Evaluate the SVD of $\boldsymbol{S_{bwn}} = \boldsymbol{V}\Sigma\boldsymbol{V}^{\mathrm{T}}$
Step 9	Define the transform matrix: $\boldsymbol{\mathcal{T}} = \boldsymbol{V}\Sigma^{-1/2}$
Step 10	Evaluate the transformed data: $\boldsymbol{\mathcal{X}_1'} = \boldsymbol{\mathcal{T}}^{\mathrm{T}}\boldsymbol{\mathcal{Z}_1}$, $\boldsymbol{\mathcal{X}_2'} = \boldsymbol{\mathcal{T}}^{\mathrm{T}}\boldsymbol{\mathcal{Z}_2}$, and $\boldsymbol{\mathcal{X}_3'} = \boldsymbol{\mathcal{T}}^{\mathrm{T}}\boldsymbol{\mathcal{Z}_3}$,
Step 11	Using the maximum length of the fused feature vector r according to Eq.(1), apply the Multiset DCA on two feature sets at a time as shown in Fig. (3).

We used multiple feature sets (M-DCA) for fusing the feature extracted from three RGB, RGB-D and RGB-SkelMHI input streams in this work. We assumed that the samples in the data matrix have been collected from c distinct action classes and m set of features, $\mathcal{X}_i \in \mathbb{R}^{i\times n}, i = 1,2,3 \ldots\ldots\ldots m$, which are selected by their rank, i.e. $rank(\mathcal{X}_1) \geq rank(\mathcal{X}_2) \ldots\ldots\ldots \geq rank(\mathcal{X}_m)$. The length of fused feature vector r are selected as given as:

$$r \leq \min\left(c-1, rank\left(\mathcal{X}_i\right), rank\left(\mathcal{X}_j\right)\right) \tag{1}$$

Figure 3 shows an example M-DCA structure for three RGB, RGB-D and Skel-MHI feature sets with $rank(\mathcal{X}_1)$, $rank(\mathcal{X}_2)$, and $rank(\mathcal{X}_3)$ respectively for three modalities. For

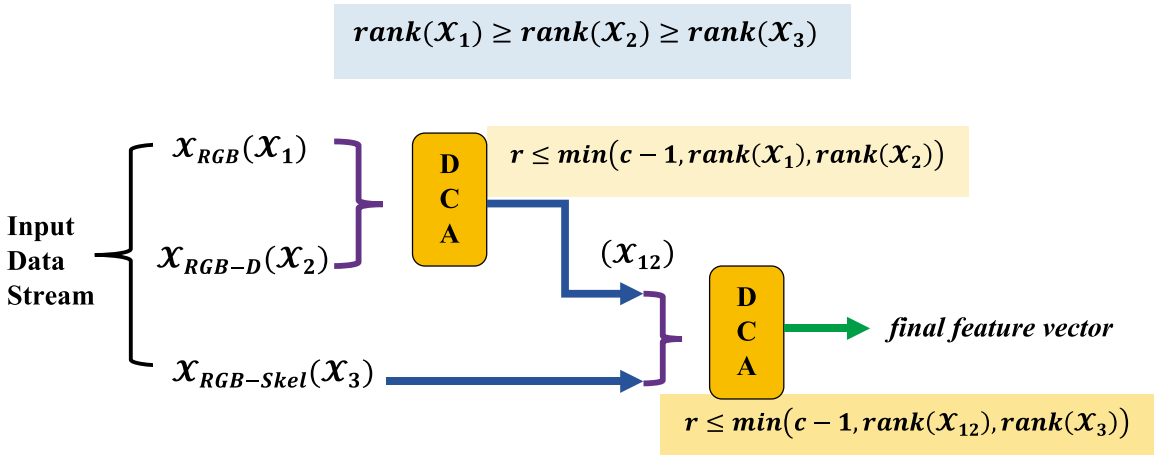


Fig. 3 Shown example of multiset DCA fusion for data streams: RGB, RGB-D and RGB-SkelMHI feature sets

instance, let's assume the $rank(\mathcal{X}_1) \geq rank(\mathcal{X}_2) \geq rank(\mathcal{X}_3)$. We fuse between $(\mathcal{X}_1)$ and $(\mathcal{X}_2)$, which have the highest ranks, in the first stage of M-DCA. It is estimated that the length of the $(\mathcal{X}_{12})$ would be greater than the length of the $(\mathcal{X}_3)$. Therefore, $(\mathcal{X}_3)$ is fused with $(\mathcal{X}_{12})$ in the next step. In this manner, for the fused feature vector, we maintain the maximum possible length in every step. In the final step, $r \leq min\left(c-1, rank\left(\mathcal{X}_{12}\right), rank\left(\mathcal{X}_3\right)\right)$ the predicted, probably shorter, feature vector length can be calculated. If the function sets of equal ranks exist, they can be combined at any time. The proposed framework merged the features obtained from three different inputs with the deep architecture at 'Conv 7b' bottleneck layer before the fully connected layers. In addition, the feature vectors derived from three data streams are fused using the Multiset DCA fusion technique at (FC) layers. Using multi-class SVM techniques, the final fused feature vectors are labelled. From the classification results described in Tables (1–4), it is observed that the representation ability increases further at lower computational complexity.

5 Experimental details and result analysis

The multimodal deep fusion framework has shown excellent classification results compared to previous approaches to human action datasets. Detailed information on the evaluation protocol for each dataset is explained in the following subsection.

5.1 Datasets

To evaluate the efficiency of the proposed system, the evaluation is carried out over four publicly available RGB-D datasets, UT Kinect, CAD 60, Florence 3D and SBU Interaction. In order to ensure transparency of results, we adopt the same evaluation protocol for each dataset as for the original work. The Leave one out cross-validation (LOOCV) training protocol used for UT Kinect, CAD 60, and Florence 3D action datasets. And the 5-fold CV evaluation scheme used for SBU Interaction dataset. The average recognition accuracy (ARA) is calculated per class over several classes $(\mathcal{C}_i)$ as per Eq. (2):

$$ARA = \frac{1}{k}\sum_{i=1}^{k}\frac{tp_i + tn_i}{tp_i + fn_i + fp_i + tn_i} \tag{2}$$

where, tp_i true positive of $(\mathcal{C}_i)$, tn_i true negative of $(\mathcal{C}_i)$, fp_i false positive of $(\mathcal{C}_i)$, fn_i false negative of $(\mathcal{C}_i)$.

5.1.1 UT kinect action dataset

This data collection is presented by Xia et al. [46] and collected by a single depth sensor with three different modalities color, depth and 3D skeleton coordinates. It consists of 10 actors performing 10 action class: waving(UT1), sitting down(UT2), throwing(UT3), sitting up(UT4), carry(UT5), picking up(UT6), walking(UT7), pulling(UT8), clapping hands(UT9), and pushing(UT10). There are 200 video sequences of a total of 6220 frames. The spatial resolution for RGB and depth maps is 640×480 and 320×240 pixels respectively and is recorded with 30 frames per second. Skeleton details are captured with 20 skeleton joints. It is a challenging dataset in terms of intra-similar activities. The UT dataset example frames are seen in Fig. 4a.

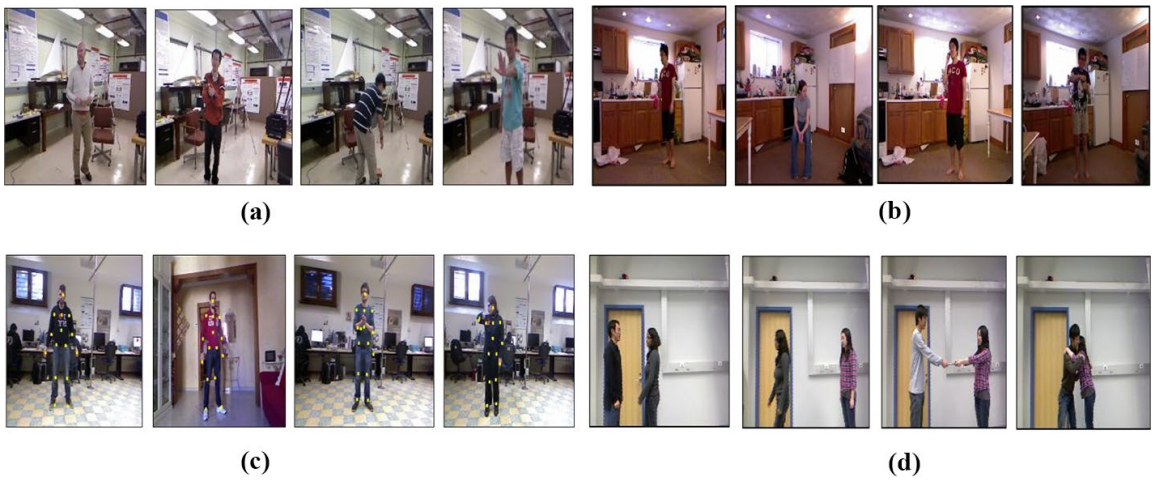


Fig. 4 Shown the example frames from **a**) UT Kinect, **b**) Florence 3D, **c**) CAD 60, **d**) SBU Interaction datasets

5.1.2 CAD 60 action dataset

Sung et al. [39] presented this dataset to identify human behavior in the different indoor home environments of living room, bathroom, office, kitchen and bedroom. It is obtained from Kinect sensor with RGB, Depth, and 3D co-ordinates in three modalities. There are four actors, 2 males and 2 females performing 12 different activities: working on computer (CA1), cooking (chopping) (CA2), rinsing mouth (CA3), relaxing on couch (CA4), opening pill container (CA5), writing on whiteboard (CA6), talking on phone (CA7), wearing contact lens (CA8), drinking water (CA9), cooking (stirring) (CA10), brushing teeth (CA11), and talking on couch (CA12). RGB and depth frames with a spatial resolution of 640x480 and skeleton joints include 15 body joint information data. The example frames from CAD 60 dataset is seen in Fig. 4b.

5.1.3 Florence 3D action dataset

This dataset is proposed by Seidenari et al. [31] and captured by the Kinect Sensor with three modalities. Ten actors are performing 9 human activities: stand up (FL1), read watch (FL2), bow (FL3), tight lace (FL4), wave (FL5), answer phone (FL6), clap (FL7), sit down (FL8), and drink from a bottle (FL9). Each actor repeated above activities two to three times. There are a total of 215 activity video sequences. The example frames from this dataset is seen in Fig. 4c.

5.1.4 SBU interaction dataset

It is proposed by Yun et al. [48] that consists of two-person activities in an indoor laboratory environment. It has been captured depth sensor with RGB video, depth frames and 3D skeleton coordinate data. It consists of seven actors that are performing eight interaction activities: hugging (SB1), punching (SB2), shaking hands (SB3), approaching (SB4), exchanging objects (SB5), departing(SB6), kicking (SB7), and pushing (SB8). This dataset is challenging because of similar body movements as well as non-periodic action classes. There is total of 300 interaction activities video composed into 21 sets. The RGB and depth frames having a

spatial resolution of 640×480 pixels with 15 fps. RGB and depth frame with a spatial resolution of 640×480 pixels with 15 fps. The 3D skeleton information is stored at 15 joints for each frame movement. The example frames from SBU dataset are depicted in Fig. 4d.

5.2 Result analysis

In this work, we extracted the features by fine-tuning the deep IR-v2 model and integrating the extracted feature vector from three different data streams at the bottleneck layer ('Conv 7b') just before the fully linked (FC) layers. It is found that the fine tuning of the 'Conv 7b' layer produced better results than the extraction of features from a fully connected layer in combination with a multi-class SVM classifier. This is because SVM achieved a better classification for intermediate layers compared to a feature derived from FC layers for small sample datasets. the recognition accuracy for each activity in four different datasets is mentioned in Tables 1–4. Furthermore, classification outcomes are shown in the confusion matrix for each dataset in Figs. 5, and 6, A comparison of the fusion accuracy obtained in percentage by different pre-trained models on the various RGB-D datasets are shown in Fig. 7.

The classification results of each activity in UT Kinect dataset is seen in Table 1, and confusion matrix is shown in Fig. 5a. It is shown that most of the dataset activities are identified with higher confidence by the proposed D-BMFF. There is slight confusion between the 'throwing' and 'carry' activities due to the occlusion caused by the human-object interaction and field of view in which captured by the sensor. However, the proposed model able to recognize all activities with the highest accuracy as compared with other existing solutions.

The classification results for each activity are seen in the CAD 60 dataset in Table 2, and confusion matrix is seen in Fig. 5b. The activities present in CAD dataset are challenging because most of them are recorded with upper body part movement. This dataset action video shows high sematic level activities that could be considered as an advantage such as drinking, eating. It can be seen from confusion matrix there is slight confusion for activities such as 'opening pill container', and 'drinking water'. The similar confusion is observed between the activities 'talking on the phone' and 'relaxing on the couch'. The remaining activities are identified with a high level of confidence by the proposed D-BMFF.

The classification results are described for each activity in the Florence 3D dataset in Table 3. The Florence 3D dataset confusion matrix is seen in Fig. 6a. For acts like 'sit down',' drinking water' and 'answering the phone', the main misunderstanding occurs. The similarity of actions and interaction between human objects performed by the actors in different ways is the primary cause for confusion. The classification results are reported for each activity in the SBU dataset in Table 4. The confusion matrix for the SBU dataset is shown in Fig. 6b. The dataset is difficult because, unlike the single actor, the interaction activities performed by the pair of actors are independent. Actions in motions are non-periodic and similar. We can see that the main misclassification is observed in the 'shaking hands' and 'exchange object' or 'approaching and departing' activities of this dataset. Instead of extremely similar behavior, the proposed D-BMFF outperformed the majority of the activities in this dataset.

Table 1 Activity wise recognition accuracy on UT Kinect dataset

Activity	**UT1**	**UT2**	**UT3**	**UT4**	**UT5**	**UT6**	**UT7**	**UT8**	**UT9**	**UT10**
Accuracy(%)	100	99	97	99	97	100	100	99	99	100

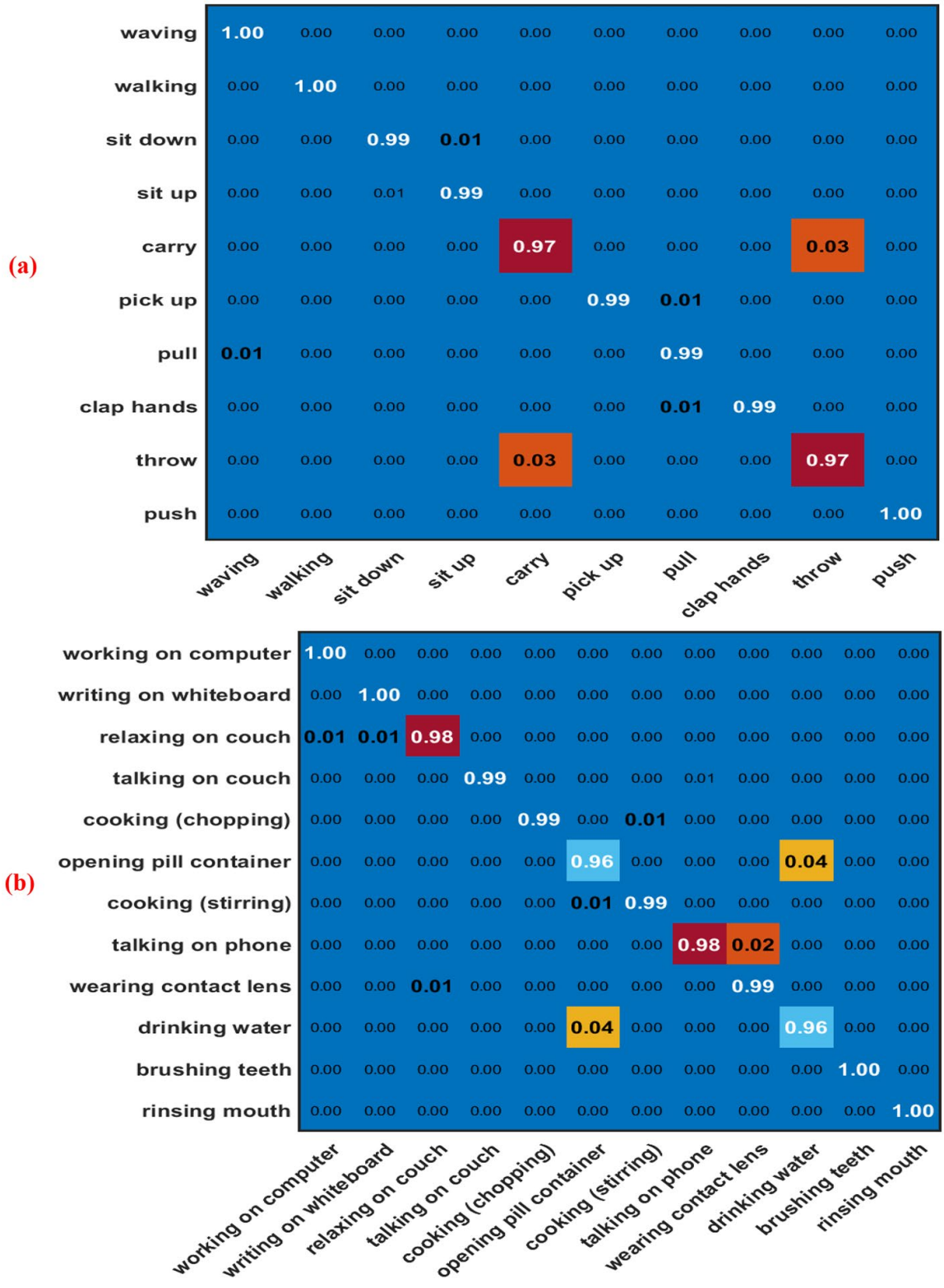


Fig. 5 Shown the confusion matrix of proposed D-BMFF technique on datasets: **a**) UT Kinect, and **b**) CAD 60

5.3 Computational time

We have evaluated the proposed framework for each dataset on a GPU machine having two NVIDIA Titan RTX 24 Gb graphic card with 128GB RAM Xeon processor, 2.20 GHz processor frequency on MatConvNet toolbox. The RGB and depth videos with a batch size

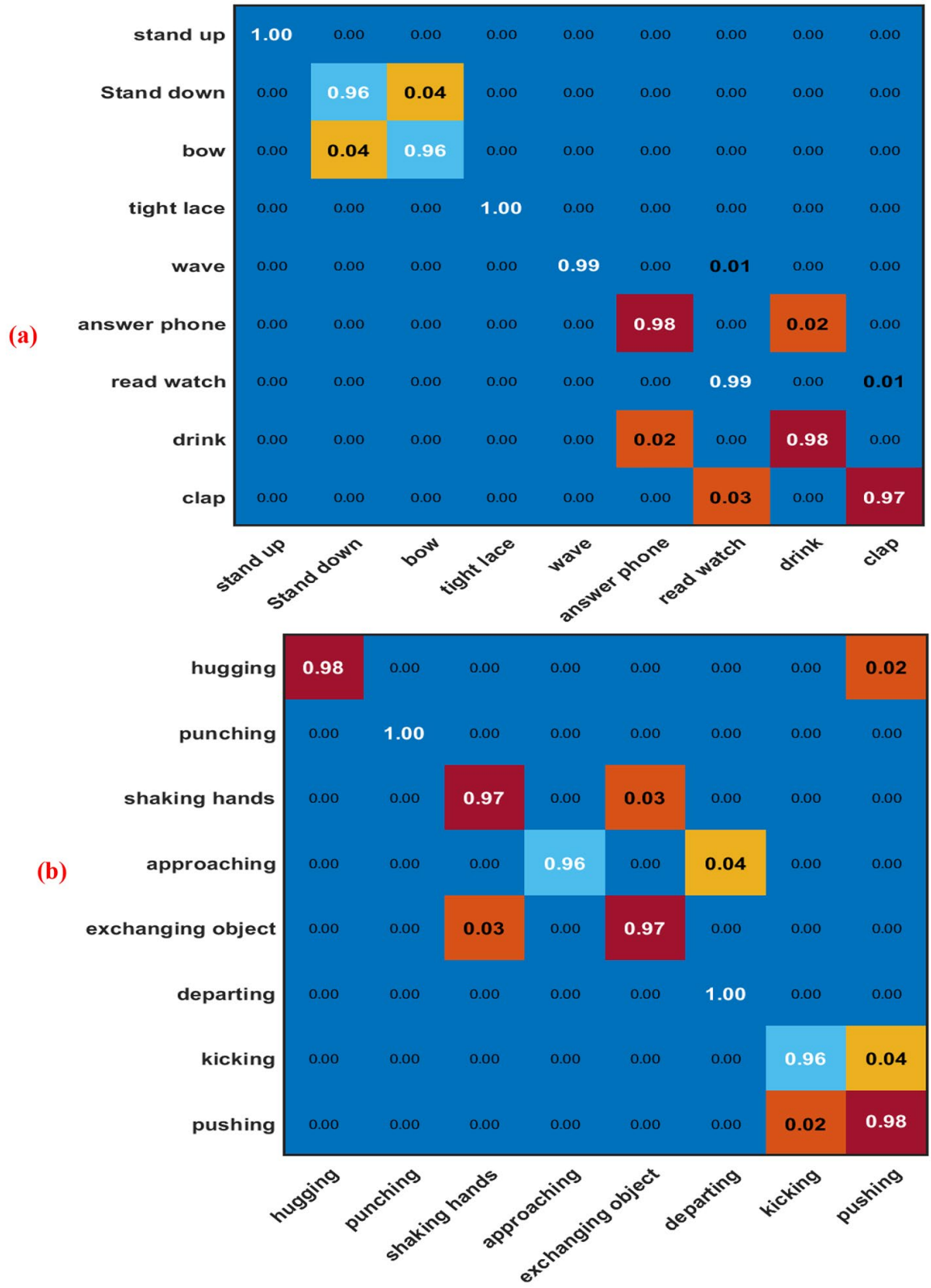


Fig. 6 Shown the confusion matrix of proposed D-BMFF technique on datasets: **a**) Florence Action 3D, and **b**) SBU Interaction

of 8 videos are fine-tuned on Conv '7b' up to 140 epochs with a dropout of 0.8 after average pooling layer to avoid the overfitting. Initially, we used the momentum constant of 0.9 and a learning rate of 0.0001. The scaling factor is choosing between 0.1 to 0.3 for residual

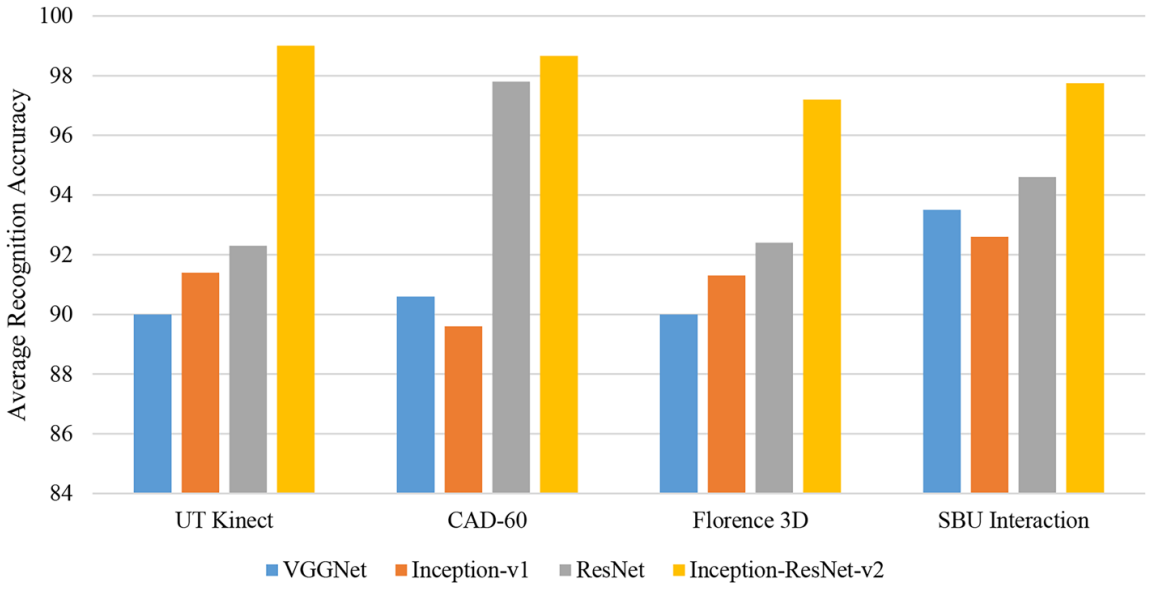


Fig. 7 Shown the fusion accuracies comparison obtained from different pre-trained models on RGB-D datasets

connection to stabilize the training data. The total running time and compared with latest approaches for each dataset is reported in Table 5. The computation time for UT Kinect dataset is 336s which includes 290.10s for training and 45.90s. The total running time for CAD-60 is 280.90s which includes 229.70s for training and 51.20 for testing. Similarly, the total computation time for Florence 3D action and SBU Interaction dataset are 349.20s and 366.70s respectively. The total running time for each dataset varies since the average training time depends on the average length of the video and total number of clips in the dataset. The state-of-the-art comparison has shown that the proposed model requires less testing time for each dataset that makes it ideal for real-time applications.

6 Multimodal deep feature fusion analysis

In this section, we discussed the results of action recognition achieved through various deep architectures. Average Recognition Accuracy (ARR) is determined based on individual data streams and multimodal fusion streams using deep models. Recognition findings from deep architecture transfer learning for UT Kinect, CAD-60, Florence Action and SBU Interaction datasets are reported in Tables 6, 7, 8 and 9. It can be shown that the proposed D-BMFF system has

outperformed each dataset. The multimodal bottleneck layer features fusion in combination with the SVM classifier greatly improves the accuracy of classification. In the D-BMFF technique, the feature vector is extracted by fine-tuning the pre-trained IR-v2 model. The three data streams of RGB, RGB(D) and Skel-MHI are fused using M-DCA at the bottleneck layer ('Conv 7b') and (FC) layers. In addition, for each action dataset,

Table 2 Activity wise recognition accuracy on CAD-60 dataset

Activity	**CA1**	**CA2**	**CA3**	**CA4**	**CA5**	**CA6**	**CA7**	**CA8**	**CA9**	**CA10**	**CA11**	**CA12**
Accuracy (%)	100	99	100	98	96	100	98	99	96	99	100	99

Table 3 Activity wise recognition accuracy on Florence Action 3D dataset

Activity	FL1	FL2	FL3	FL4	FL5	FL6	FL7	FL8	FL9
Accuracy (%)	100	99	96	100	99	98	97	96	98

the proposed approach is tuned to different deep architectures such as Inception-v1 [40], ResNet [13] , and VGGNet [35].

In order to know the contribution of each modality to action representation, we performed an experiment using the individual data streams RGB, RGB(D) and Skel-MHI, and the data are tabulated in Tables 6–9. The result of the different deep pre-trained models on UT Kinect dataset is listed in Table 6. The recognition accuracy of feature fusion obtained at Conv '7b' layer is 6% higher than 'FC' layer by using IR-v2 architecture. Table 6 shows that the individual data stream classification accuracy is less than the combination of a feature fusion. This is because knowledge about shape and motion can be seen to be correlated in depth sequences, but these are hard to record simultaneously. In addition, the 3D skeleton joint coordinates do not differentiate against such actions because of noise and self-occlusion. For example, some activities 'eating' and 'drinking' having the same temporal pattern that cannot be identified using joint coordinates.

The feature fusion results are increased by a significant margin when the feature are extracted from intermediate Conv '7b' layer as compared with the fully connected layer. The feature fusion accuracy at bottleneck layer is sufficiently increased for CAD-60 (8%,), Florence 3D (6%), and SBU (7%) datasets reported in Tables 7–9. It is noticed that skeleton streams given the maximum accuracy for each dataset as compared with RGB and depth frames. This is because 3D coordinates are first converted into motion history skeleton images for each activity videos. The RGB-SkelMHIs incorporates temporal pose variations and robust to variations due to similar action classes.

Table 10 indicated the accuracy of the classification with the combination of RGB, RGB(D) and SkelMHI on each dataset. The deep features are extracted from pre-trained IR-v2 model for each data stream combination. It can be seen that the intermediate layer feature fusion worked much better than the fully connected layer. This illustrated the feasibility of the multimodal fusion method by inadequately exploiting the robust data held by multiple input modalities. The increase in accuracy also showed that RGB and depth data are complementary to each other (as well as RGB and SkelMHI).

In the next section, the classification result by the proposed D-BMFF framework on each dataset is compared with state-of-the-art approaches.

7 Comparison with state-of-the-art

It is experimentally found that feature fusion at bottleneck layer before the top layer performed well in conjunction with the SVM classifier. The ARR on UT Kinect, CAD-60, Florence 3D, SBU Interaction are 99%, 98.50%, 98.10%, and 97.75% respectively. The

Table 4 Activity wise recognition accuracy on SBU Interaction dataset

Activity	SB1	SB2	SB3	SB4	SB5	SB6	SB7	SB8
Accuracy (%)	98	100	97	96	97	100	96	98

Table 5 Comparison of computational run on human activity datasets

Dataset Works	UT Kinect	CAD-60	Florence 3D	SBU Interaction
She et al. [33]	190.30s(Train) 72.90s(Test)	-	-	294.93 (Train) 117.86 (Test)
Li et al. [24]	-	59.37s (Train) 129.80s(Test)	-	-
Proposed	290.10s(Train) 45.90s(Test)	229.70s (Train) 51.20s(Test)	295.50s(Train) 53.70s(Test)	310.90s(Train) 55.80s(Test)

recognition results showed the efficacy of D-BMFF framework. We achieved the highest accuracy on CAD-60, Florence 3D and SBU interaction dataset and comparable performance on UT Kinect dataset. The recognition results by the different works on UT Kinect dataset is mentioned in Table 11. It is seen from Table 11, most of the works utilized only skeleton input for action representation [3, 26]. The proposed accuracy on UT dataset is comparable with the work [26] as they fused the depth and skeleton information. However, it seems easy to implement the model using skeleton information only, but they suffer due to self-occlusion and noise errors to estimates the skeleton coordinates. Our method showed higher recognition rate than the works [8, 11, 49] by a sufficient margin. In addition, multimodal deep feature fusion using three data stream shown better results as a compared combination of other both streams.

Table 12 depicted the state-of-the-art accuracy comparison on CAD-60 dataset. It is noted that the proposed model showed the highest accuracy as compared with similar multimodal features based works [20, 21, 29]. The recognition rate is higher in our work because we utilized color Skel-MHI images for feature extraction that is robust to noise and illumination variations.

The results obtained on Florence 3D dataset are shown in Table 13. It is observed that proposed D-BMFF showed higher accuracy despite the intra-class similarity and different way of performing the action classes. The multimodal feature fusion (M-DCA) at intermediate layer performed better way as compared with works [21, 29].

The state-of-the-art comparison of SBU interaction datasets is seen in Table 14. The proposed model outperforms the other existing approaches to this challenging dataset. We achieved the highest accuracy of 97.75% on such complex interaction activity dataset. The D-BMFF gained a sufficient margin as compared with only skeleton features based activity descriptors [9, 48].

Table 6 Result comparison of deep architectures on UT Kinect dataset

Model Dataset	Modality	VGGNet [35] (FC)	Inception-v1 [40] (FC)	ResNet [13] (FC)	Inception-ResNet-v2	
					(FC)	(Conv '7b')
UT Kinect	RGB	84.50	80.70	85.40	86.00	90.70
	Depth	82.50	85.10	86.20	87.70	91.20
	Skel-MHI	86.40	88.50	89.40	86.60	90.10
	Fusion	90.00	91.40	92.30	93.60	**99.00**

Table 7 Result comparison of deep architectures on CAD-60 dataset

Model Dataset	Modality	VGG19 [35] (FC)	Inception-v1 [40] (FC)	ResNet [13] (FC)	Inception ResNet-v2	
					(FC)	(Conv '7b')
CAD 60	RGB	85.80	87.00	86.10	89.30	94.40
	Depth	84.40	88.70	89.60	85.70	91.30
	Skel-MHI	89.10	88.50	91.10	89.80	95.60
	Fusion	90.60	89.60	97.80	90.10	**98.66**

Table 8 Result comparison of deep architectures on Florence 3D dataset

Model Dataset	Modality	VGG19 [35] (FC)	Inception-v1 [40] (FC)	ResNet [13] (FC)	Inception ResNet-v2	
					(FC)	(Conv '7b')
Florence 3D	RGB	86.50	85.60	87.70	90.60	92.90
	Depth	84.60	88.10	89.60	87.40	91.20
	Skel-MHI	86.70	92.30	90.50	88.20	93.40
	Fusion	90.00	91.30	92.40	91.35	**97.20**

Table 9 Result comparison of deep architectures on SBU Interaction dataset

Model Dataset	Modality	VGG19 [35] (FC)	Inception-v1 [40] (FC)	ResNet [13] (FC)	Inception ResNet-v2	
					(FC)	(Conv '7b')
SBU Interaction	RGB	90.80	90.00	92.80	92.40	96.80
	Depth	89.50	89.80	90.10	89.30	92.30
	Skel-MHI	91.40	90.50	92.10	87.70	95.60
	Fusion	93.50	92.60	94.60	90.70	**97.75**

Table 10 Accuracy comparison of the combination of RGB, RGB(D), and SkelMHI on Action Dataset

Modality Dataset	**RGB+RGB(D)**	**RGB(D)+SkelMHI**	**RGB+ RGB(D)+SkelMHI**	**M-DCA Fusion**
UT Kinect	85.50	92.60	93.60	(FC)
	91.80	96.40	**99.00**	(Conv '7b')
CAD 60	84.00	91.50	90.10	(FC)
	92.20	94.00	**98.66**	(Conv '7b')
Florence 3D	87.00	89.50	91.35	(FC)
	90.40	94.30	**97.20**	(Conv '7b')
SBU Interaction	91.80	90.60	90.70	(FC)
	94.50	95.00	**97.75**	(Conv '7b')

Table 11 Comparison of multimodal features fusion score using Inception-Resnet-v2 on UT Kinect dataset

Works	Input	Method	Accuracy (%)
Avola et al. [3]	RGB + Depth	BoW + ELM	84.00
Devanne et al. [8]	Skeleton	Riemannian Manifold+ kNN	91.50
Ghodsi et al. [11]	Skeleton	Joint skeleton trajectories	96.80
Zhang et al. [49]	Skeleton	Multilayer LSTM	97.00
Vemulapalli et al. [42]	Skeleton	Lie Group	97.08
Huynh-The et al. [14]	Skeleton	PAM + Pose-Transition	97.00
Nguyen et al. [26]	Depth +skeleton	KM-IELLogE-IELLogE	99.50
Proposed	RGB+ Depth	**D-BMFF**	91.80
	Depth +skeleton		96.40
	RGB+ Depth+ Skeleton		**99.00**

Table 12 Comparison of multimodal features fusion score using Inception-Resnet-v2 on CAD 60 dataset

Works	Input	Method	Accuracy(%)
Avola et al. [3]	RGB+ Depth	BovW +ELM	82.60
Sung et al. [39]	Depth +skeleton	Max Entropy Markov Model	83.20
Raman and Maybank [29]	Depth+ Skeleton	H-HMM	85.40
Khaire et al. [20]	RGB+ Depth+ Skeleton	5-CNNs	90.00
Kong et al. [21]	RGB+ Depth+ Skeleton	CMFL	94.10
Liu et al. [25]	Skeleton	Pose+ kNN	95.75
Li et al. [23]	Skeleton	ShapeDTW	97.30
Franco et al. [10]	RGB+ Skeleton	HOG+BOW fusion	98.30
Proposed	RGB+ Depth	**D-BMFF**	92.20
	Depth +skeleton		94.00
	RGB+ Depth + Skeleton		**98.50**

Table 13 Comparison of multimodal features fusion score using Inception-Resnet-v2 on Florence 3D Action dataset

Works	Input	Method	Accuracy (%)
Seidenari et al. [31]	Skeleton	NBNN Bag-of-Poses	82.00
Devanne et al. [8]	Skeleton	Riemannian Manifold	87.04
Vemulapalli et al. [42]	Skeleton	Lie Group	90.88
Salih and Youssef [30]	Skeleton	STIP MSH	86.13
Huynh-The et al. [14]	Skeleton	PAM + Pose-Transition	92.09
Yang et al. [47]	Skeleton	Latent Max-Margin	93.42
Nguyen et al. [26]	Skeleton +Depth	KM-IELLogE-IELLogE	95.37
Proposed	RGB+ Depth	D-BMFF	90.40
	Depth +skeleton		94.30
	RGB+ Depth+ Skeleton		**98.10**

Table 14 Comparison of multimodal features fusion score using Inception-Resnet-v2 on SBU Interaction dataset

Works	Input	Methods	Accuracy (%)
Yun et al. [48]	Skeleton	SVM and MIL Boost	87.30
Feng et al. [9]	Skeleton	Bi-Vector and LSTM	82.70
Ijjina and Chalavadi [15]	RGB + Depth	ConvNet + ELM Classifier	86.58
Keçeli et al. [19]	Depth	2D and 3D ConvNet	94.70
Khaire et al. [20]	RGB + Depth + Skeleton	VGG-F+WPM	96.26
Proposed	RGB+ Depth	**D-BMFF**	94.50
	Depth +skeleton		95.00
	RGB+ Depth+ Skeleton		**97.75**

8 Conclusion

Deep learning-based architectures are dominant and widely used for computer vision applications, especially for visual recognition task. Since single modality dependent approaches are severely affected by self-occlusions, illumination variation, view-variation, and noise present in video sequences. In this work, a deep bottleneck multimodal feature fusion (D-BMFF) human activity recognition system is proposed that utilize three data modalities RGB, Depth, and 3D Skeleton data. It is observed that DCA is a compelling feature fusion technique that maximizes pairwise correlations between the features set and restricts the correlation to be within classes with low computational complexity. The multimodal feature fusion increases the recognition rate and allows to make the best use of the available data. Experimentally, it is discovered that the feature fusion at the bottleneck layer before the top layer worked well in conjunction with the SVM classifier. In order to assess the efficacy of the proposed model, we performed an experiment on four standard human activity datasets. Recognition results showed that the proposed framework outperformed other state-of-the-art methods to all these datasets. We have achieved the best accuracy of the CAD-60, Florence 3D and SBU interaction dataset and equivalent performance of the UT Kinect dataset.

However, the extraction of multimodal features from intermediate layers is more complicated due to a large number of deep layers of pre-formed architectures. In the future, we will try to develop a deep model using parallel computing and representation of the CNN architecture. We will also expand our approach to applications in real-time, for elderly care and with more complex actions datasets.

References

1. Aggarwal J, Xia L (2013) Human activity recognition from 3D data: A review. Pattern Recognit Lett 48
2. Amor BB, Su J, Srivastava A (2016) Action Recognition Using Rate-Invariant Analysis of Skeletal Shape Trajectories. IEEE Trans Pattern Anal Mach Intell 38(1):1–13
3. Avola D, Bernardi M, Foresti GL (2019) Fusing depth and colour information for human action recognition. Multimed Tools Appl 78:5919–5939
4. Blank M, Gorelick L, Shechtman E, Irani M, Basri R (2005) Actions as space-time shapes. In 10th IEEE Int Conf Comput Vis (ICCV'05) Beijing

5. Bobick AF, Davis JW (2001) The recognition of human movement using temporal templates. IEEE Trans Pattern Anal Mach Intell 23(3):257–267
6. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297
7. Dalal N, Triggs B, Schmid C (2006) Human detection using oriented histograms of flow and appearance. In Proc Eur Conf Comput Vis
8. Devanne M, Wannous H, Berretti S, Pala PD, Bimbo AD (2015) 3-D Human Action Recognition by Shape Analysis of Motion Trajectories on Riemannian Manifold. IEEE Trans Cybern 45(7):1340–1352
9. Feng J, Zhang S, Xiao J (2017) Explorations of skeleton features for LSTM-based action recognition. Multimed Tools Appl 1–13
10. Franco A, Magnani A, Maio D (2020) A multimodal approach for human activity recognition based on skeleton and RGB data. Pattern Recognit Lett 131:293–299
11. Ghodsi S, Mohammadzade H, Korki E (2018) Simultaneous joint and object trajectory templates for human activity recognition from 3-D data. J Vis Commun Image Ranging 55:729–741
12. Haghighat M, Mottaleb MA, Alhalabi W (2016) Discriminant Correlation Analysis: Real-Time Feature Level Fusion for Multimodal Biometric Recognition. IEEE Trans Inf Foren Secur 11(9):1984–1996
13. He K, Zhang X, Ren S, Sun J (2015) Deep residual learning for image recognition. arXiv preprint arXiv:1512.03385
14. Huynh T-T, Hua C-H, Tu NA, Hur T, Bang J, Kim D, Amin MB, Kang BH, Seung H, Shin S-Y, Kim E-S, Lee S (2018) Hierarchical topic modeling with pose-transition feature for action recognition using 3D skeleton data. Inf Sci 444:20–35
15. Ijjina EP, Chalavadi KM (2017) Human action recognition in RGB-D videos using motion sequence information and deep learning. Pattern Recognit 72:504–516
16. Ji S, Xu W, Yang M, Yu K (2013) 3D convolutional neural networks for human action recognition. IEEE Trans Pattern Anal Mach Intell 35(1):221–231
17. Ji X, Cheng J, Feng W, Tao D (2017) Skeleton embedded motion body partition for human action recognition using depth sequences. Signal Process 143:56–68
18. Ji Y, Yang Y, Xu X, Shen HT (2018) One-shot learning based pattern transition map for action early recognition. Signal Process 143:364–370
19. Keçeli AS, Kaya A, Can AB (2018) Combining 2D and 3D deep models for action recognition with depth information. Signal Image Vid Process 12:1197–1205
20. Khaire P, Kumar P, Imran J (2018) Combining CNN streams of RGB-D and skeletal data for human activity recognition. Pattern Recognit Lett 000:1–10
21. Kong J, Liu T, Jiang M (2019) Collaborative multimodal feature learning for RGB-D action recognition. J Vis Commun Image Ranging 59:537–549
22. Laptev I (2005) On Space-Time Interest Points. Int J Comput Vis 64(2–3):107–123
23. Li Q, Lin W, Li J (2018) Human activity recognition using dynamic representation and matching of skeleton feature sequences from RGB-D images. Signal Process Image Commun 68:265–272
24. Li R, Fu H, Lo W-L, Chi Z, Song Z, Wen D (2019) Skeleton-Based Action Recognition With Key-Segment Descriptor and Temporal Step Matrix Model. IEEE Access 7
25. Liu T, Wang J, Hutchinson S, Meng MQ-H (2019) Skeleton-Based Human Action Recognition by Pose Specificity and Weighted Voting. Int J Soc Robot 11:219–234
26. Nguyen XS, Mouaddib AI, Nguyen TP (2019) Hierarchical Gaussian descriptor based on local pooling for action recognition. Mach Vis Appl 30:321–343
27. Phyo CN, Zin TT, Tin P (2019) Deep Learning for Recognizing Human Activities Using Motions of Skeletal Joints. IEEE Trans Consum Electron 65(2):243–252
28. Prabhakar S, Jain A (2002) Decision-level fusion in fingerprint verification. Pattern Recognit 35(4):861–874
29. Raman N, Maybank S (2016) Activity recognition using a supervised non-parametric hierarchical HMM. Neurocomputing 19:163–177
30. Salih AAA, Youssef C (2016) Spatiotemporal representation of 3D skeleton joints-based action recognition using modified spherical harmonics. Pattern Recognit Lett 83:32–41
31. Seidenari L, Varano V, Berretti S, Bimbo A, Pala P (2013) Recognizing Actions from Depth Cameras as Weakly Aligned Multi-Part Bag-of-Poses. In 3rd International Workshop on Human Activity Understanding from 3D data, Portland, Oregon
32. Shabaninia E, Nilchi ARN, Kasaei S (2019) A weighting scheme for mining key skeletal joints for human action recognition. Multimed Tools Appl 78:31319–31345
33. She Q, Mu G, Gan H, Fan Y (2020) Spatio-temporal SRU with global context-aware attention for 3D human action recognition. Multimed Tools Appl 79:12349–12371
34. Simonyan K, Zisserman A (2014a) Two-stream convolutional networks for action recognition in videos. In Proc Adv Neural Inf Process Syst

35. Simonyan K, Zisserman A (2014b) Very deep convolutional networks for large-scale image recognition. arXiv:arXiv:1409.1556
36. Singh T, Vishwakarma DK (2020) A Deeply Coupled ConvNet for Human Activity Recognition using Dynamic and RGB Images. Neural Comput Appl
37. Singh T, Vishwakarma DK (2018) Video benchmarks of human action datasets: a review. Artif Intell Rev 52(2):1107–1154
38. Sun Q-S, Zeng S-G, Liu Y, Heng P-A, Xia D-S (2005) A new method of feature fusion and its application in image recognition. Pattern Recognit 38(12):2437–2448
39. Sung J, Ponce C, Selman B, Saxena A (2012) Unstructured human activity detection from RGBD images. In IEEE Int Conf Robot Autom Saint Paul MN
40. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In IEEE Conf Comput Vis Pattern Recognit
41. Szegedy C, Ioffe S, Vanhoucke V, Alemi A (2016) Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. arXiv:1602.07261v2 [cs.CV]
42. Vemulapalli R, Arrate F, Chellappa R (2014) Human Action Recognition by Representing 3D Skeletons as Points in a Lie Group. In IEEE Conf CVPR Columbus OH
43. Vishwakarma DK, Singh T (2019) A visual cognizance based multi-resolution descriptor for human action recognition using key pose. AEU Int J Electron Commun 107:157–169
44. Wang H, Schmid C (2013) Action recognition with improved trajectories. In Proc Int Conf Comput Vis (ICCV)
45. Wang H, Klaeser A, Schmid C, Liu CL (2013) Dense trajectories and motion boundary descriptors for action recognition. IJCV
46. Xia L, Chen C, Aggarwal J (2012) View Invariant Human Action Recognition Using Histograms of 3D Joints. In Comput Vis Pattern Recognit Workshops (CVPRW)
47. Yang Y, Deng C, Tao D, Zhang S, Liu W, Gao X (2017) Latent Max-Margin Multitask Learning With Skelets for 3-D Action Recognition. IEEE Trans Cybern 47(2):439–448
48. Yun K, Honorio J, Chattopadhyay D, Berg TL, Samaras D (2012) Two-person Interaction Detection Using Body-Pose Features and Multiple Instance Learning. In Comput Vis Pattern Recognit Workshops (CVPRW) Rhode Island
49. Zhang S, Liu X, Xiao J (2017) On Geometric Features for Skeleton-Based Action Recognition Using Multilayer LSTM Networks. In IEEE Winter Conf Appl Comput Vis (WACV) Santa Rosa CA

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

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

A study on transformer-based Object Detection

Hritik Vaidwan
Dept. of Applied Physics
Delhi Technological University
New Delhi, India
hritikvaidwan@gmail.com

Nikhil Seth
Dept. of Applied Physics
Delhi Technological University
New Delhi, India
nikhilseth.08@gmail.com

Anil Singh Parihar
Dept. of Computer Science and Engineering
Delhi Technological University
New Delhi, India
parihar.anil@gmail.com

Kavinder Singh
Dept. of Computer Science and Engineering
Delhi Technological University
New Delhi, India
Kavinder85@gmail.com

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498550

***Abstract*— This paper focuses on transformers based end-to-end object detection methods. End to end object detection is a new paradigm that has got attention in recent times. It does not require complex hand-engineered components such as non-max suppression to detect objects inside an image. Various methods are proposed to date to enhance fully end-to-end object detectors, most of them are based on the attention mechanism. In this work, we analyze some algorithms which involve transformers for the purpose of object detection. We discuss end-to-end models in which we have focused on Adaptive clustering-based transformers which solve attention encoder redundancy, Deformable Detection Transformers in which the attention module attends a limited collection of key sampling points, Unsupervisedly pre-trained Detection Transformers which are pre-trained on random query patches from the given image to improve accuracy and finally the Transformer-based Set Prediction using FCOS. These enhanced models not only improve the mean average precision of the model but also improves the total convergence time.**

Keywords*— *End to end object detection; Transformer; Attention; DETR; ACT; UP-DETR; TSP-FCOS

I. INTRODUCTION

Object detection techniques using deep neural networks have evolved significantly with time. It all started with Convolution neural networks (CNN) that significantly boosted the performance of computer vision, image processing and many other processes. Lots of advancements are achieved on CNN to come up with several object detection architectures such as RCNN [1], SPP-Net [2], Fast-RCNN [3], Faster-RCNN [4], YOLO [5]–[8], etc. Researchers proposed transformers based End-to-End object detection method (DETR) which introduced a completely new paradigm for object detection using an altered version of the transformer. Object detection techniques employ many hand-craft components such as anchor generation [9], rule-based training target assignment, non-maximum suppression (NMS) [10], post-processing and many more which are not completely end-to-end. DETR eliminates the need for such common hand-crafted components, and built the first completely end-to-end object detector, achieving very competitive performance.

II. RELATED WORK

Before starting to study end-to-end object detection techniques [37], we should know "what a Transformer is" and how it utilizes "self-attention" that makes it stand out? So that we can easily acknowledge the concept of end-to-end object detection.

A. Feature extraction of images

Detecting objects in an image is a difficult task, it is required for numerous applications such as face detection in security cameras [38], autonomously driving vehicles [39], medical disease detection such as cancer cells [40], etc. Earlier, object detection models use Haar-Like features [41], HOG [42], and other methods for extracting features from an image. In CNNs, the machine handles the task of feature extraction, which helps in achieving better performance as compared to traditional methods. The convolutional neural networks (CNNs) act as the base of various object detection methods. That's why they are also referred to as backbone in structures to extract features from the image. Several networks of CNNs are commonly used in the object detection task, for example, ResNet [13], AlexNet [11] and VGGNet [12]. These networks are generally used for accurate object classification tasks and have been evaluated on some widely used large datasets and benchmarks such as ImageNet [15], COCO [14], etc. In object detection methods, various versions of ResNet are used. A unique feature in ResNet models is the introduction of the so-called "identity shortcut connection", which skips one or more layers, which led to better depth in the network leading to better feature extraction.

B. Transformers

The Encoder-Decoder model uses several recurrent units like RNN, GRU, LSTM, etc. Each recurrent unit in the encoder-decoder processes the information in sequence by taking input of previous hidden state & input in sequence and generates an output. The last hidden state generated by the encoder is assumed to have all relevant information of sequence and is passed to the decoder. Decoder recurrent units similarly perform computation taking input of previous hidden state & generating output.

This model is quite powerful because it can easily map sequences of different lengths to each other. However, it performs poorly on complex tasks since the single encoding vector is not capable of holding information about the whole sequence. To overcome them, Transformers were introduced, it eliminates recurrence & it uses an attention mechanism to come up with relations between fed input &

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:50:57 UTC from IEEE Xplore. Restrictions apply.

output. Self-attention relates different positions of a given single sequence to compute a representation of the same sequence. The transformer uses point-wise fully connected layers & stacked self-attention layers in both encoder & decoder.

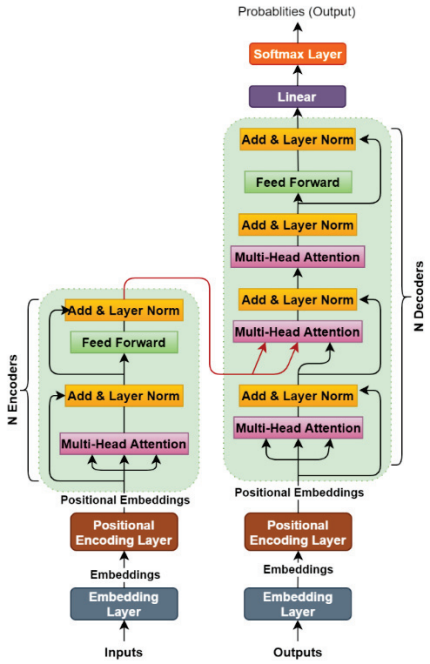


Fig. 1. Architecture of a transformer

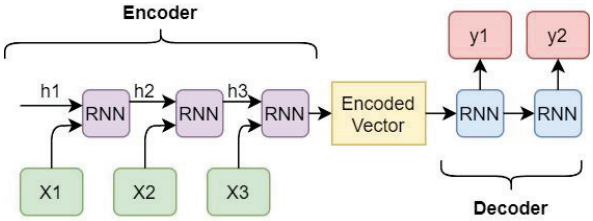


Fig. 2. Encoder-Decoder Sequence to Sequence model architecture.

The encoder comprises N stacked layers, each having multi-head attention [16] module & feed-forward network [17]. The decoder also comprises N stacked layers, each layer has a multi-head attention module, masked multi-head attention & feed-forward network which makes the autoregressive nature of the Transformer. Due to the non-recurrent nature of transformers, they are fed with positional encodings [16] to facilitate sequence-like structure. The main advantages of transformers over other models are their global computation & perfect information-carrying capacity which makes them ideal than RNNs.

III. Methodology

Firstly, the image is fed into a CNN network usually called a backbone. This backbone extracts out a feature vector from the image. This generated feature vector is now fed as an input to the Transformer. The feature vector is converted to the desired dimension to make it in a sequence. Finally, the output from the Transformer is fed to an FFN (Feedforward network) which predicts the bounding boxes & various classes. FFNs can be retrained based on the type of requirement.

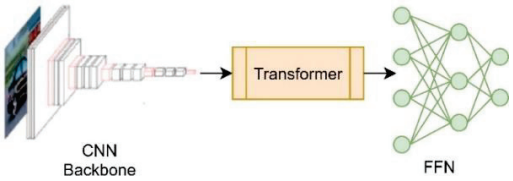


Fig. 3. Basic Structure of End-to-End detectors. (FFN represents Feed Forward Neural Network)

A. DETR – DEtection TRansformer

In DETR [1], Carion *et al.* proposed an encoder-decoder architecture based on Transformers for the object detection task. As discussed in the "Attention is all you need" [2], the self-attention [2] operation, simply takes a weighted average of all pairs between each element of a given sequence, thus makes them a highly suitable option for specific constraints of a set-prediction problem such as removing duplicated predictions. The main features of DETR are 1) Integration of a new bipartite matching loss [3] function that scrapes the need for non-max suppression in the object detection task, and 2) modifying the Language transformers with a parallel decoding technique that makes a suitable trade-off between global computations required for set prediction & computational cost. DETR takes a fixed-size vector of N predictions, called object queries as input in a parallel fashion in the decoder, where N is chosen to be significantly larger than the total number of objects in an image.

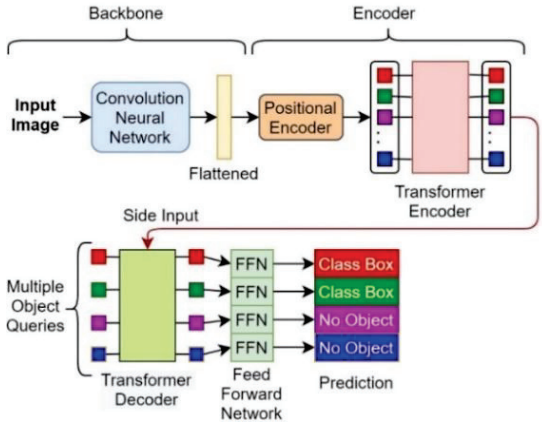


Fig. 4. Architecture of DETR

The overall DETR architecture is made of 3 components:

a CNN backbone [20], an encoder-decoder architecture & FFN [17]. An image is fed into the backbone & a feature vector is given as output, typically of size C x H x W. Further, this is passed through a 1x1 convolution [21] to reduce the channel dimension from C to a smaller dimension d. The encoder expects a sequence of inputs, so the vector is flattened to dx(HW) size. Since the transformer has a permutation invariant architecture, a fixed sine based encoding technique is used and these encodings are fed to the input of each layer in the encoder. Decoder transforms N embeddings of size d in parallel. Here the decoder is also not affected by permutation, so the N input embeddings to the decoder must be different to produce different results. These embeddings are also called object queries and these are learnt positional embeddings.

The output encodings given by the decoder are fed to a three-layer perceptron having ReLU [22] activations in FFNs independently which are converted to box coordinates &

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:50:57 UTC from IEEE Xplore. Restrictions apply.

class labels. The model globally reasons about all objects together using pairwise relations between them using encoder-decoder attention method & self-attention over these embeddings, while being able to use the whole image as context.

Auxiliary decoding losses are used in decoders during training, as it helps the model predict the correct number of objects in each class. Hungarian loss [23] is added after each decoder layer. Training settings differ from regular object detectors like the DETR requires extra-long schedules & uses auxiliary decoding losses in transformers.

For this study, the ResNet-50 backbone based DETR model is chosen, having 6 encoder layers, 6 decoders layers. The model has 41.3 million parameters & achieves on avg. 40.6 on short schedules & 42.0 AP on long schedules respectively. A test on encoding layers confirms that AP drops by 3.9 points & 6.0 on large objects without these encoding layers. According to the hypothesis concluded, the encoder is essential for object disentanglement. The encoder divides instances, which should make the decoder's object extraction and localization process easier. The FFNs inside transformers are important in the process, as removing them leads to a reduction in performance by 2.3 AP. An analysis is done by Carion *et al.* on positional encodings in both encoder & decoder show that learnt positional encodings did not affect the result much at the encoder, however, learned encodings at the output of the decoder, were better when learned at attention than at input.

B. Deformable DETR

In Deformable DETR [24], Zhu et al. describes that DETR [18] performs poorly on detecting small objects in an image. It is because other object detectors usually work with multi-scale features, where small objects are detected using high-resolution feature maps. But in DETR, high-resolution feature maps lead to various complexities and make it even more computationally expensive and hence they are not preferred. Zhu et al. proposed a Deformable attention module that only attends a limited collection of key sampling points around a reference point, regardless of the spatial scale of the function maps. When only a small fixed number of keys are assigned for each query, the issues of convergence and feature spatial resolution can be mitigated. In the decoder, the deformable attention module is designed for processing convolutional feature maps as key elements, it is inspired by the deformable convolution technique [25]. It can be noted that top-down structure in FPN [26] is no more used as it takes a lot of time to train, rather a newly proposed multi-scaled deformable attention technique is used as it has the ability in itself to exchange information between multi-scaled feature maps.

In the decoder, there are cross attention and self-attention modules. The cross attention module has been replaced with a multi-scale deformable attention module as the deformable attention module is designed for processing convolution feature maps as key elements. Hence, the self-attention module is still unchanged in deformable DETR. For each object query, the 2D normalized coordinate of the given reference point is predicted from its object query embedding through a learnable linear projection and then it is passed to a sigmoid [27] function. In this technique, the attention module is used to extract features around the given reference point, that lets the detection head predict the bounding box as relative offsets concerning the given reference point so that the optimization problem can further be reduced.

One more improvement that Zhu et al. introduced to the deformable DETR is that it uses multi-scale deformable attention. This is almost similar to the older single-scale version, except that it samples more points from multi-scale feature maps. Older techniques focus only on one sampling point for each attention head. However, multi-scale deformable attention looks over multiple sampling points from multi-scale inputs. This technique can also be perceived as an efficient variant of transformer attention, where a pre-filtering mechanism is introduced by the deformable sampling locations. When the sampling points traverse all possible locations, the proposed attention module is equivalent to Transformer attention.

Some other minor improvements over DETR [18] presented by Zhu et al. are: a new effective iterative bounding box refinement mechanism is used to increase detection performance. This technique is better as here each decoder layer refines bounding boxes based on the prediction it gets from the previous decoder layer. Moreover, a two-stage Deformable DETR is used here, where generated region proposals are fed into the decoder as object queries for refinement. For region proposal generation, the decoder is removed here and the encoder only Deformable DETR is used. Inside it, each pixel is assigned as an object query, which is used to predict bounding boxes. Top scoring bounding boxes are then selected as region proposals. No Non-Max-Suppression [10] applied before feeding the region proposals to the second stage.

C. UP transformers (Unsupervised Pre-training for Object Detection with Transformers)

CNN backbone in DETR with around 23.2M parameters is pre-trained to give good visual representations, but the transformer structure has not been pre-trained and hence the model has the possibility of improvement. Liu et al. in "Unsupervised Pre-Training for object Detection with transformers" [28] came out with a solution. The technique is inspired by unsupervised pre-training used in natural language processing [29] which improves the accuracy of models to a great extent. UP-DETR [28] makes the DETR model converge faster and manages to provide better accuracy on datasets like PASCAL VOC [37] and COCO [14].

In this approach, multiple query patches from the given image are randomly cropped & then pre-training of the complete model takes place to predict bounding boxes of these randomly cropped query patches. The model is pre-trained for the detection of these query patches from the original image which are not annotated. But this attracts some problems, to avoid query patch detection to destroy the classification features, the backbone and patch feature reconstruction are frozen to preserve feature discrimination of the transformers during the training process. Also, to solve any possible assignment problem between query patches and object queries, an object query shuffle and attention mask [30] is introduced. The ultimate process of Unsupervised pre-training follows 2 steps: (1) Pre-training on a large-scale dataset with the pretext task, (2) finetune the parameters on downstream tasks.

Firstly, in this method, a frozen backbone (CNN) is used in the model to draw out a visual representation with the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:50:57 UTC from IEEE Xplore. Restrictions apply.

feature map of an input image, where channel dimension is denoted by C & H x W is the feature map size. After that, positional encodings are added with the feature map and passed to the multi-layer transformer encoder. For the randomly cropped query patches, the CNN backbone with global average pooling (GAP) [31] is used to extract the patch feature, which is then flattened and supplemented with object queries and then finally passed into a transformer decoder. Object query points to the position embeddings fed to the decoder. At the time of pre-training, the model takes random query patches as input and then its decoder predicts the bounding boxes corresponding to the position of those patches. In the process of cropping random query patches, the model crops the query patches with a random coordinate, width, height and are resized to 128 x 128 pixels, transformed with the SimCLR-style [32] without using horizontal flipping, random colour distortion, & Gaussian blurring techniques. Liu et al. presented that freezing pre-trained CNN is important to pretrain transformers. Besides, it also validates the pretext (random query patch detection) may weaken the feature discrimination of the pre-training CNN, & localization and classification have different feature preferences.

UP-DETR achieves faster convergence at the early training stage with feature reconstruction. With frozen CNN and feature reconstruction, UP-DETR heavily relies on a pre-trained CNN model. Therefore, a good direction is to further investigate UP-DETR with random query patch detection and contrastive learning together to pretrain the whole DETR model from scratch.

D. Adaptive Clustering Transformer (ACT)

Zheng et al. [33] proposed an adaptive Clustering Transformer (ACT) which is based on the drawbacks of high computational requirements of DETR for high-resolution input. It clusters the query features adaptively using a technique called Locality Sensitive Hashing (LSH) [34] and approximates the query-key interaction using the prototype key interaction. Zheng et al. replaced the self-attention module on the original pre-trained DETR-model, without much decrease in its performance. Due to its compatibility with the original Transformer, it does not require retraining. The performance gap due to the substitution of ACT can be further reduced by equipping with Multitask Knowledge Distillation (MTKD) which also facilitates seamless switching between models with different FLOPs & accuracy during training. Two observations of DETR motivate ACT:

1) Encoder Attention Redundancy: Inside the encoder of DETR, features at each position assimilate information from other spatial positions by attention mechanism. Similar features, nearer to each other, will generate similar attention maps, vice versa. This redundancy in self-attention maps, motivates ACT to choose representative prototypes & broadcast the feature update of prototypes to its nearest neighbours.

2) Encoder Feature Diversity: As the encoder layers increase, features start becoming similar to each other & collect information from each other. This motivates ACT to adaptively determine the no. of prototypes based on the distribution of features in each layer.

To solve attention encoder redundancy, ACT clusters similar queries & calculates key-query attentions only for representative prototypes. These features are then broadcasted to the nearest neighbour according to Euclidean distance. Encoder Feature diversity leads to the designing of an algorithm that can cluster features according to the distribution of them in the respective encoder layer of each transformer. Thus, a multi-round Exact Euclidean Locality Sensitivity Hashing(E2LSH) [34] performs query feature distribution-aware clustering.

ACT can reduce the quadratic $O(n^2)$ complexity inside self-attention to O(nk), where k is no of prototypes finalized in each layer.

E. TSP-FCOS

In "Rethinking Transformer-based Set Prediction for Object Detection" [36], Sun et al. stated that the Bipartite matching Hungarian loss & the cross-attention mechanism of the transformer have some instability problems, resulting in a slower convergence. Sun et al. came out with a few improvements to fix the slower convergence problem: A new Feature of Interest (FoI) selection mechanism is introduced in TSP-FCOS [36] architecture so that the transformer encoder can work with multi-level features, as a result, this model outperforms in the terms of object detection accuracy. Also, a new bipartite matching scheme is designed for the model for accelerating the training process which helps the model to converge in less than 50 epochs. Feature extraction is done from the input image using a CNN backbone. It builds the Feature Pyramid Network (FPN) [26] component using the feature map from the backbone, which outputs multi-level features that help encoder-only DETR to detect objects of various sizes. The multi-level features are then distributed to two subnets, one of which is the classification subnet (Head), which performs FoI classification on them & the other is an auxiliary subnet. The encoder then receives the concatenated outputs from them, which are then picked by the FoI classifier before being sent to the encoder.

Following that, Self-attention [16] is used to aggregate the information from different FoIs, and the encoder outputs are passed through an FFN, which predicts the bounding box and class label (including "no class") for each FoI.

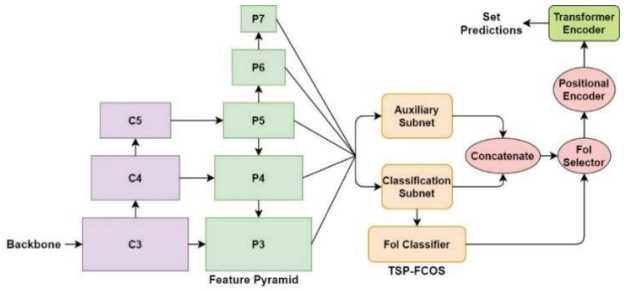


Fig. 5. TSP-FCOS Object detection network architecture.
(Here C3, C4 and C5 are the feature maps added to the backbone and P3, P4, P5, P6, P7 are feature maps added to the Feature pyramid network (FPN)).

The Transformer's positional encoding is generalized to a 2D image scenario following DETR. The positional encoding of feature points with a normalized position is defined as [PE(x): PE(y)], here PE is a function defined as:

$$PE(x)_{2i} = \sin(x / 10000^{2i/d_{model}}) \quad (1)$$

$$PE(x)_{2x+1} = \cos(x / 10000^{2i/d_{model}}) \quad (2)$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:50:57 UTC from IEEE Xplore. Restrictions apply.

where d_{model} is the dimension of the Feature of Interest(s).

To compare with DETR, a training strategy which is similar to DETR is used by *Sun et al*., where a 96-epoch (8x) training schedule and random crop augmentation is applied and after that, it has been found that TSP-FCOS significantly outperform DETR [18].

IV. Comparision

In this section, we will discuss the contrasting features between the methodologies which use the end-to-end object detection paradigm. In the table, we will see the method and its salient features. Table I shows comprehensive study:

TABLE I. METHODS AND THEIR FEATURES

METHOD	FEATURES
DETR	• Vanilla Transformer • Uses Direct set Prediction Problems. • High computations requirements
Deformable DETR	• Multi-scaled deformable attention modules are introduced, which attend to a small set of key sampling points near a reference point. • Top-down structure in FPN is not used, as multi-scale deformable attention makes them redundant. • Can even detect small objects with good accuracy.
Adaptive Clustering Transformer	• Clustering of queries, to reduce the identical attention maps generated. • E2LSH is used for mapping identical queries. • A tradeoff between performance & speed is made. • MTKD is used for seamless switching between DETR & ACT.
UP DETR	• Training of transformers happens unsupervisedly pre-trained on a large-scale dataset without annotations. • Converges in lesser training time. • Global average pooling is used in the backbone (GAP) • The entire model is fine-tuned with labelled data on the downstream tasks.
TSP-FCOS	• bipartite matching scheme is designed for the model for accelerating the training process. • Feature of Interest (FoI) selection mechanism is introduced so that the transformer encoder can work with multi-level features.

Inference time refers to the time taken by a trained machine learning algorithm to make a prediction. Inference time is a vital factor for deciding the speed and practicality of the real-time machine learning prediction model for obvious reasons as the lesser the inference time, the lesser will be the lag in the real-time applications. FLOPS stands for "Floating Point Operations Per Second," and it is a common indicator of computing power (i.e., how fast a machine runs). A billion FLOPS is referred to as a GigaFlop or Gflop. It is a more precise measurement than counting the number of instructions per second. FLOPs is used as a unit by machine learning model to calculate how much computation power is needed for some task. mAP or Mean Average Precision is a precision metric used in object detection problems. In this, a precision-recall curve of the given dataset is plotted after deciding a particular IoU threshold and then the area under the curve is measured to find mAP.

In the COCO-2017 challenge, to find mAP, the average of AP of All object classes is taken with 10 IoU thresholds. In the PASCAL-VOC-2007 challenge, AP for each object is measured with IoU threshold of 0.5. Similarly, AP for all objects is calculated and averaged to find mAP. Generally, a ResNet backbone is trained on a COCO training set and then it is evaluated using the COCO validation set with various score cut-offs such as AP50, AP75, etc. In practical applications, it is used to train the convolution layers of the object detection techniques as it helps the model to learn basic shapes of the objects in the initial convolution layers, after that the trained neural network gets fine-tuned as per the model's output requirements.

Table II shows comparative analysis based on GFLOPS and Average Precision.

TABLE II. COMPARISON OF MODELS

Model	GFLOPS	AP	AP_S	AP_M	AP_L
Faster RCNN-DC5	320	39.0	21.4	43.5	52.5
Faster RCNN-FPN	180	40.2	24.2	43.5	52.0
Faster RCNN-R101-FPN	246	42.0	25.2	45.6	54.6
DETR	**86**	**42.0**	**20.5**	**45.8**	**61.1**
DETR-DC5	**187**	**43.3**	**22.5**	**47.3**	**61.1**
DETR -R101	152	43.5	21.9	48.0	61.8
DETR-DC5-R101	253	44.9	23.7	49.5	62.5
UP-DETR-R50		**42.8**	**20.8**	**47.1**	**61.7**
TSP-FCOS	255	**44.4**	**27.7**	**48.6**	57.3
Deformable DETR	173	**43.8**	**26.4**	**47.1**	58.0
Deformable DETR -Two-Stage	173	46.2	28.8	49.2	61.7
ACT-R50 -DC5(L=32)	168.9	**42.6**	**21.4**	**46.8**	61.1
ACT-R50-DC5 -MTKD (L=32)	168.9	43.1	22.2	47.1	61.4
ACT-R50 -DC5(L=16)	155.7	37.1	13.9	41.3	58.8
ACT-R50-DC5 -MTKD (L=16)	155.7	40.6	18.5	44.3	59.7

V. CONCLUSION

In this paper, we have presented a study on few end-to-end object detection methods which are based on transformer mechanism. This study shows various challenges and limitations of the existing algorithms. We noticed that the modifications to the attention layers or the backbone of

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:50:57 UTC from IEEE Xplore. Restrictions apply.

transformers may result in substantial performance improvements, as seen in the preceding discussion.

Transformers can be unsupervisedly pre-trained on a large-scale dataset without annotations to converge in a very less number of epochs and this helped UP-DETR to secure the highest AP on a large-sized image dataset. Similarly, ACT can be used without much training & reduces the computation and memory costs for object detection. In TSP-FCOS, A Feature of Interest selection mechanism enables transformers-encoder to work with multi-level features, where the top-scored features can be chosen for better accuracy, it is proven to be highly accurate on a medium-sized image dataset and the remaining one is Deformable DETR, that has an excellent adaptation on small-sized images due to its deformable attention mechanism which helps it to attain the highest AP on small-sized image dataset.

Despite the substantial improvement of models in recent times, some unanswered problems, such as better convergence, less training & computation remain to be addressed in the future research.

REFERENCES

[1] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 580–587, 2014, doi: 10.1109/CVPR.2014.81.

[2] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 8691 LNCS, no. PART 3, pp. 346–361, 2014.

[3] R. Girshick, "Fast R-CNN," Proc. IEEE Int. Conf. Comput. Vis., vol. 2015 Inter, pp. 1440–1448, 2015.

[4] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137–1149, 2017.

[5] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," Proc. IEEE Conf. Comput. Vis. Pattern Recognit., vol. 2016-Decem, pp. 779–788, 2016.

[6] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017, vol. 2017-Janua, pp. 6517–6525, 2017.

[7] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," arXiv, 2018.

[8] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," arXiv, 2020.

[9] Y. Zhong, J. Wang, J. Peng, and L. Zhang, "Anchor box optimization for object detection," Proc. - 2020 IEEE Winter Conf. Appl. Comput. Vision, WACV 2020, pp. 1275–1283, 2020.

[10] J. Hosang, R. Benenson, and B. Schiele, "Learning non-maximum suppression," Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017, vol. 2017-Janua, pp. 6469–6477, 2017.

[11] M. Z. Alom et al., "The history began from AlexNet: A comprehensive survey on deep learning approaches," arXiv, 2018.

[12] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc., pp. 1–14, 2015.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2016-Decem, pp. 770–778, 2016.

[14] T. Y. Lin et al., "Microsoft COCO: Common objects in context," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 8693 LNCS, no. PART 5, pp. 740–755, 2014.

[15] O. Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge," Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, 2015.

[16] A. Vaswani et al., "Attention is all you need," Adv. Neural Inf. Process. Syst., vol. 2017-Decem, no. Nips, pp. 5999–6009, 2017.

[17] D. Dai, W. Tan, and H. Zhan, "Understanding the feedforward artificial neural network model from the perspective of network flow," arXiv, 2017.

[18] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-End object detection with transformers," arXiv, 2020.

[19] D. Georgiev and P. Liò, "Neural Bipartite Matching," arXiv, 2020.

[20] Q. Zhang, M. Zhang, T. Chen, Z. Sun, Y. Ma, and B. Yu, "Recent advances in convolutional neural network acceleration," Neurocomputing, vol. 323, pp. 37–51, 2019.

[21] C. Szegedy et al., "Going deeper with convolutions," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 07-12-June, pp. 1–9, 2015.

[22] A. F. M. Agarap, "Deep Learning using Rectified Linear Units (ReLU)," arXiv, no. 1, pp. 2–8, 2018.

[23] R. Stewart, M. Andriluka, and A. Y. Ng, "End-to-End People Detection in Crowded Scenes," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2016-Decem, pp. 2325–2333, 2016.

[24] X. Zhu, W. Su, L. Lu, B. Li, X. Wang, and J. Dai, "D EFORMABLE DETR : D EFORMABLE T RANSFORMERS FOR END - TO -E ND O BJECT D ETECTION," 2017.

[25] J. Dai et al., "Deformable Convolutional Networks," Proc. IEEE Int. Conf. Comput. Vis., vol. 2017-Octob, pp. 764–773, 2017.

[26] Z. Zhong, L. Sun, and Q. Huo, "An anchor-free region proposal network for Faster R-CNN-based text detection approaches," Int. J. Doc. Anal. Recognit., vol. 22, no. 3, pp. 315–327, 2019.

[27] C. E. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation functions: Comparison of trends in practice and research for deep learning," arXiv, pp. 1–20, 2018.

[28] Zhigang Dai, Bolun Cai, Yugeng Lin, and Junying Chen. Up-detr: Unsupervised pre-training for object detection with transformers. arXiv preprint arXiv:2011.09094, 2020.

[29] Y. Liu and Z. Lin, "Unsupervised Pre-training for Natural Language Generation: A Literature Review," arXiv, pp. 1–10, 2019.

[30] M. Kimura and M. Tanaka, "Intentional attention mask transformation for robust CNN classification," arXiv, pp. 1–9, 2019.

[31] M. Lin, Q. Chen, and S. Yan, "Network in network," 2nd Int. Conf. Learn. Represent. ICLR 2014 - Conf. Track Proc., pp. 1–10, 2014.

[32] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," arXiv, no. Figure 1, 2020.

[33] M. Zheng, P. Gao, X. Wang, H. Li, and H. Dong, "End-to-End Object Detection with Adaptive Clustering Transformer," 2020, [Online]. Available: http://arxiv.org/abs/2011.09315.

[34] O. Jafari and P. Nagarka, "Experimental Analysis of Locality Sensitive Hashing Techniques for High-Dimensional Approximate Nearest Neighbor Searches," arXiv, 2020.

[35] Sun, Zhiqing & Cao, Shengcao & Yang, Yiming & Kitani, Kris. (2020). Rethinking Transformer-based Set Prediction for Object Detection.

[36] M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The Pascal Visual Object Classes Challenge: A Retrospective," *Int. J. Comput. Vis.*, vol. 111, no. 1, pp. 98–136, 2015, doi: 10.1007/s11263-014-0733-5.

[37] AS Parihar, K Singh, Vardaan, SD Gonade, "A comprehensive analysis of recent advancements in object detection", IJAST, vol. 29, no. 7, pp. 13473 - 13484, Jul. 2020.

[38] P. Anghelescu, I. Serbanescu and S. Ionita, "Surveillance system using IP camera and face-detection algorithm," *Proceedings of the International Conference on ELECTRONICS, COMPUTERS and ARTIFICIAL INTELLIGENCE - ECAI-2013*, Pitesti, Romania, 2013, pp. 1-6, doi: 10.1109/ECAI.2013.6636188.

[39] M. V. Rajasekhar and A. K. Jaswal, "Autonomous vehicles: The future of automobiles," 2015 IEEE International Transportation Electrification Conference (ITEC), Chennai, India, 2015, pp. 1-6, doi: 10.1109/ITEC-India.2015.7386874.

[40] A. S. Parihar, "A study on brain tumor segmentation using convolution neural network," 2017 International Conference on Inventive Computing and Informatics (ICICI), Coimbatore, India, 2017, pp. 198-201, doi: 10.1109/ICICI.2017.8365336.

[41] Rezaei, Mahdi & Nafchi, Hossein & Morales, Sandino. (2013). Global Haar-Like Features: A New Extension of Classic Haar Features for Efficient Face Detection in Noisy Images. 10.1007/978-3-642-53842-1_26.

[42] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005, pp. 886-893 vol. 1, doi: 10.1109/CVPR.2005.177.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:50:57 UTC from IEEE Xplore. Restrictions apply.

ACADEMIA
Accelerating the world's research.

IJERT-A Survey on Intrinsically Motivated Reinforcement Learning

IJERT Journal

International Journal of Engineering Research and Technology (IJERT)

Cite this paper

Downloaded from Academia.edu

Get the citation in MLA, APA, or Chicago styles

Related papers

Download a PDF Pack of the best related papers



DEEP REINFORCEMENT LEARNING: AN OVERVIEW
tanoy dewanjee

IJERT-Reinforcement Learning:Recent Threads
IJERT Journal

Playing Atari with Deep Reinforcement Learning
Qinglan Peng

Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 05, May-2021

A Survey on Intrinsically Motivated Reinforcement Learning

Shweta Meena
Department of Software Engineering
Delhi Technological University
New Delhi, India

Abhay Singh Hyanki
Department of Software Engineering
Delhi Technological University
New Delhi, India

Tarun Kumar
Department of Software Engineering
Delhi Technological University
New Delhi, India

***Abstract*—Machine learning (ML) consists of mainly three further studies that are supervised learning, unsupervised learning, and reinforcement learning. In this paper, we are going to look at the later part that is reinforcement learning. We will review the various methods that have been and are being used in the field of reinforcement learning. Reinforcement learning basically deals with how software agents tend to behave to maximize its reward in a given environment. Rewards are also of two types: extrinsic and intrinsic. Extrinsic reward is some specific outcome that we achieve after following certain procedures and completing a given task. In contrast, the intrinsic reward is the curiosity of the agent to develop new skills that might be useful in the coming future.**

Keywords*— *Intrinsic Rewards, Extrinsic Rewards, Reinforcement learning

I. INTRODUCTION

Reinforcement learning has developed its place in ML and AI communities within last 7-9 years. It offers to build a way of programming subjects by the medium of rewards for successful trials and punishment for wrong attempts without telling the agent the procedure to complete the given task .But these processes face various problems. In this paper we look at those problems and also look for solutions. Machine learning have three sub-parts that are supervised learning, unsupervised learning, and reinforcement learning. All three parts are depicted in Fig 1.

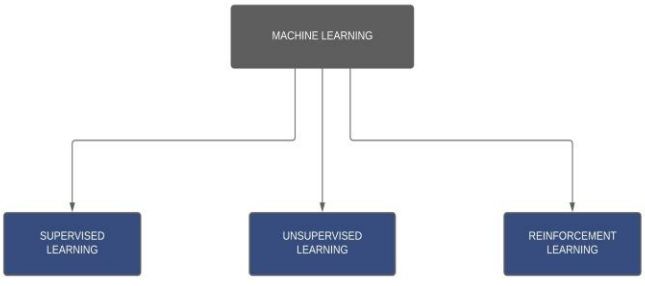


Fig. 1 Sub-parts of machine learning

A. *Supervised Learning*

Supervised learning is the machine learning technique in which we develop a function which best fits the training dataset. Depending on the training with the dataset, the machine is able to predict future outputs given a random input. It uses a function that is gathered from available labeled training data, which consists of some sets of various training examples. In supervised learning, all the examples consists an input object that is also called vector and an output value also called as supervisory signal. The algorithm that is used in supervised learning examines the given training data and develops an inferred function, that is generally used for making new examples. Figure 2 below depicts the basic structure of supervised learning.

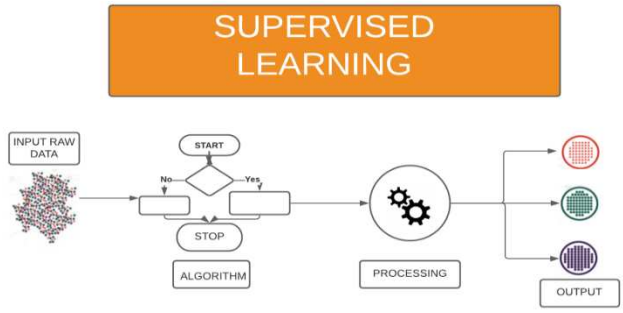


Fig. 2 Structure of Supervised learning

B. *Unsupervised Learning*

Unsupervised learning is a study in which it looks for previous undetected patterns in a given data set which does not contain any pre-existing labels and with no or very less human interaction or supervision. In comparison with supervised learning, which is known for using human-labeled data, the unsupervised learning, which can also be called as self-organization, permits users for required modeling of the probability densities over the inputs. The structure of unsupervised learning is shown in Fig 3.

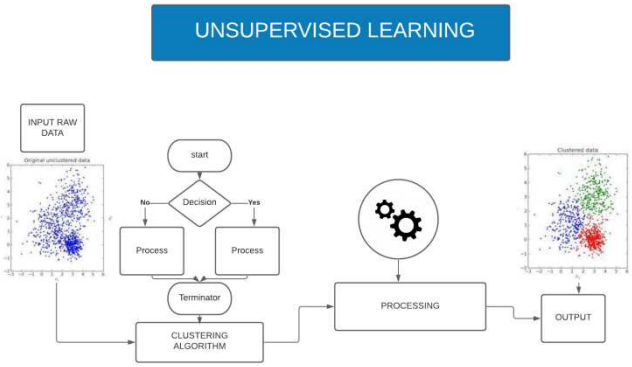


Fig. 3 Structure of Unsupervised learning

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 05, May-2021

C. *Reinforcement Learning*

Reinforcement learning (RL) is considered to be an area that is concerned with how agents or software agents perform actions in a given environment with the main motive of maximizing the cumulative reward [1].

RL is different from supervised learning as it does not requires labeled input data or output pairs to be available, it also differs as it does not requires various sub-optimal actions that are expected to be accurate or corrected. The focus of RL is to maintain an equilibrium between exploration (of the uncharted territory) and exploitation (of the current knowledge)

The algorithms for RL in this context use dynamic programming techniques. Hence, the Markov decision process is used for stating the environment. Reinforcement learning differs from classic dynamic programming algorithms because reinforcement learning usually does not assume an understanding of the exact mathematical model of the Markov Decision Process,[2] and they target large MDPs where the same methods become infeasible. The methodology used in general reinforcement learning is depicted in the below diagram Fig 4.

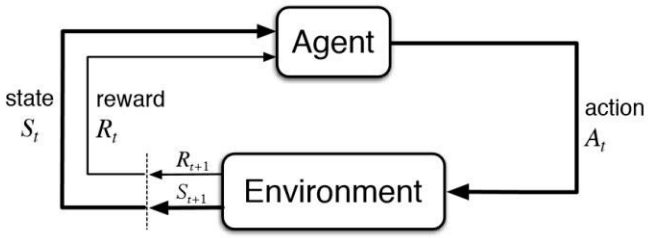


Fig. 4 Reinforcement learning

There have been many breakthroughs in reinforcement learning. Many games such as Pacman, Minigrid, Atari games, etc[3] are developed on the concept of reinforcement learning. In reinforcement learning, the agent tends to maximize its reward. In this paper, we are going to look deep into the field of reinforcement learning and describe the various methods used in reinforcement learning.

II. OVERVIEW OF REINFORCEMENT LEARNING

RL is termed as the training of some machine learning models which are used to make some decisions or sequence of decisions. The agent starts learning to achieve its goal in a potentially complex and uncertain environment [3]. In RL artificial intelligence has to face a situation that is very similar to a game. The computer tries to solve the problem by trial and error method. In order to achieve the task, artificial intelligence gains rewards or gets a penalty for the action it performs. The aim is to maximize the total reward [4].

The reward policy of the game or the rules of the game is set by the designer, but the designer does not give the artificial intelligence or the agent suggestions to solve the game. The agent has to figure it out by himself, and the main motive of the agent is to maximize its rewards. It basically starts with random trial and error methods, but as the game proceeds, it develops skills and starts implementing strategies [5]. Currently, we can say that reinforcement learning is the most efficient method to showcase the creativity of a machine.[6]

A. *Examples of reinforcement learning*

Reinforcement learning was restricted in the past because the infrastructure of the computers was not up to the mark. Actual development happened after the production of backgammon AI.

- Good examples of RL would be that of a car. While driving, the computer doesn't get driving instructions. The programmer will allow the machine to learn from its failures.
- We cannot predict the requirements of the car as to when it should drive slow or when it should go fast. To program this, the programmer would need to use lengthy multiple "if-then" statements. Instead of using these statements, the programmer uses a agent that can learn from its experience in which it gets rewards and penalties for its actions.
- Development of prosthetic legs is a challenging task for a coder as it would require extensive coding to detect the patterns of a walk. Hence, Stanford Neuromuscular Biomechanics Laboratory has developed a musculoskeletal model based on RL that has the ability to detect people's walking patterns.

B. *Challenges with reinforcement learning*

Setting up the environment is the main challenge in RL. It can be easy sometimes, like in the cases of Atari games, chess, etc. but when it comes to real-world problems, it can cause a lot of questions. Moreover, transferring the model into the real world from the simulation environment can get tricky. Scaling the neural network, which is controlling our agent, is another challenge. System of reward and penalties is the only way to communicate with the network, which would result in catastrophic forgetting, which means that when we gain new knowledge causes old knowledge to get erased. Another challenge is when the agent does not perform the task in the required way. For example, If we need our agent to walk straight, but it starts jumping, it can also be a big problem.

RL has two types of rewards that are intrinsic and extrinsic. Earlier, only extrinsic rewards were present, so what was the need of introducing intrinsic rewards.

Sometimes extrinsic rewards were sparse and even not present hence the introduction of intrinsic rewards was necessary. Intrinsic reward is basically the curiosity of the agent to develop new skills as the agent goes further in the game. Intrinsic motivation leads agents to explore the environment

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 05, May-2021

driven by curiosity when an explicit reward is not present. Such activities help in the development of competence instead of focusing on external goals. In contrast, generally, machine learning algorithms do not perform well with new problems when they arise after some time. The skills gained during competence are "building blocks," using which an agent possesses the ability to solve various issues as they appear in its lifetime.

We have identified four challenges in Deep Reinforcement Learning to which intrinsic motivation is the answer.

a) Sparse rewards

Major RL algorithms function in environments where there is no shortage of rewards, *i.e.,* the agent will gain a reward after completing an action. The agent will receive a reward signal after executing an immense sequence of steps in cases of an environment with sparse rewards. In a game called Montezuma's revenge, the agent has to pick up different objects while moving from one room to another; the agent gets rewards only when it picks up an object or receives the reward when it has left the room. This game is a typical example of a sparse reward function. Such environments that have sparse rewards are very difficult to solve considering the exploration policies mentioned above as the computer or agent does not get any hint in the way as to how to improve its exploration policy. Due to this, the agent fails to find its rewards concerning the given task [7]. Generally, to encounter this problem, an intermediate function called dense reward is added to the prize associated with the task rather than focusing on improving the exploration policies [8]. And due to this adding of the additional reward function, some more unexpected errors arise, and most of the time, expert knowledge is required.

b) Building a good state representation

It is said that, in RL, the good state representation has to be Markovian, representing the policy's actual value, which should be of low dimesions and be able to generalize [9]. Using these feature space in learning a given task can speed up the learning process and can even help in some other computations that can be learning an effective forward model [10]. Constructing a minimal feature space with disentangle features is the best and most effective way to do this.

Relevant state representation is significant in reinforcement learning. To understand it, we will consider a navigation task in which the agent aims to get to a specified location in a given space. Suppose the computer or agent is able to access pixels input from the empty space above. In that case, the agent will then have to determine its existing position and the required target position by using complex non-linear transformations in order to get the idea in which direction it has to move. In contrast to this [11], if the agent already have the idea or access to its existing position, then the only work of the agent will be comparing its vertical and horizontal positions and check whether they are greater than, smaller than, or equal to the target. In standard RL , this problem is worsened due to the availability of back propagation of reward signal as the only available process of learning, and also due to the availability or presence of sound or noise in its raw state. The interaction among agents is considered to be rich in information, yet the agent doesn't gain any knowledge from the interaction if the reward is less or sparse. Moreover, the state representation, which has a reward associated with it, is entirely dependant on a specific task and cannot take other tasks into account. And in contrast, the state representation, which is learned without associating with any task, can be generalized for the remaining tasks. There have been many works regarding the understanding of good representation of state.

c) Temporal abstraction of actions

High-level actions, usually called options, are used under the Temporal abstraction of actions, and they can be executed at different times [2]. At the time of execution of the option, due to the presence of an intra option policy [12], the action can be defined in each state. The number of actions executed after choosing an option is usually termed as the length of the option and is usually fixed. The options that are to be accomplished are usually selected by the inter option policy. By using options, the number of decisions that are to be taken is reduced, and abstract actions help in attaining so hence abstract actions can be considered as a key element in accelerating the learning process. The credit assignment [2]problem can also be solved by this easily. In this problem, the rewards arise or occurs with a slight delay and don't majorly affect any of the temporary distant states which have occurred before it, however these temporary distant states may be required in the process of obtaining the reward. In fact, the agent will have to carry the reward throughout the list of actions performed in order to apply the state action tuple which was involved at the very beginning. This problem also focuses on determining the importance of the action, and by performing which action the agent can get the reward[13]. If the action sequence is very large, this procedure can be very slow.

d) Building a curriculum

It usually occurs in the framework of the study known as multi-task RL in which a single agent tries to solve more than one task or several tasks [14]. It is all about describing a schedule throughout the process. It can be gathered from the observation that when examples are arranged in a specific or meaningful order [11], the learning becomes much easier. Moreover, a curriculum might arrange or organize tasks or actions in a way that they are more

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 05, May-2021

Algorithm	Merits	Demerits
SARSA	Always tries to find the best policy for exploration.	Does not contain the ability to function efficiently at local minima
DQN	It is comparatively faster in training	It estimates more Values of Q than the desired amount
Q-Learning	Explores more than one policy at a time	It is a bit slow but can handle alternative routes
Monte Carlo	Deals with unknown model risen difficulty i.e. strategy decision	After one episode the strategy can change
Value-function approximation	more stable and less prone to failure	Less sample efficient
Policy Gradient actor critic	Finds the Optimal Deterministic Policy	Typically converge to a local optimum rather than a global optimum

close to each other and can also be increasingly complex. For example, in helping the robot move an ice cube by itself, an effective curriculum would be to teach the robot how to grab the ice cube first. In this way, the robot can learn from the skills that it gained during the process of grasping the cube. Since moving an ice cube requires a vast amount of knowledge like the movement of joints, so without any prior knowledge of movements of joints, which can be learned through the first process, it would be very difficult. Making pre-specified tasks or action sequences as the curriculum is one of the standard methods or expert scores can also be used, which acts as the baseline score [15] Some other methods are also there like one method which was given by Florensa requires strong assumptions, and another method given by Wu relies on task decomposition. It tells us that, generally, in standard methods, there is a requirement of an expert in curriculum learning always.

In table 1, we have discussed the various algorithms used in the field of reinforcement learning. The table consists of name of the algorithm and their merits and demerits.

III. CONCLUSION

There have been many breakthroughs in RL. RL has been used in many games such as Pacman, Atari games, etc. RL is the most efficient way of finding out the creativity of the machine.

Hence, the study of RL is going to become more important in the near future. In this paper, we outlined the importance of reinforcement learning and the challenges that arise in reinforcement learning. However, these challenges can be tackled by intrinsic motivation as a reward. Intrinsic motivation is the next big step in reinforcement learning.

RL is being used in automation of operations, maintenance of machinery, optimization of usage of energy. Many industries like finance are also considering RL to power training systems which work on artificial intelligence. We know that the trial-and-error technique of training robots is time consuming, then also this is practised as it allows the robots in using their skills in order to evaluate the real world more efficiently and also complete the tasks using their skills.

REFERENCES

[1] P. R. Montague, “Reinforcement Learning: An Introduction, by Sutton, R.S. and Barto, A.G.,” *Trends Cogn. Sci.*, 1999, doi: 10.1016/s1364-6613(99)01331-5.

[2] R. S. Sutton and A. G. Barto, “Reinforcement Learning: An Introduction,” *IEEE Trans. Neural Networks*, 1998, doi: 10.1109/tnn.1998.712192.

[3] N. Bougie and R. Ichise, “Skill-based curiosity for intrinsically motivated reinforcement learning,” *Mach. Learn.*, 2020, doi: 10.1007/s10994-019-05845-8.

[4] A. G. Barto, “Intrinsic motivation and reinforcement learning,” in *Intrinsically Motivated Learning in Natural and Artificial Systems*, 2013.

[5] T. D. Kulkarni, K. R. Narasimhan, A. Saeedi, and J. B. Tenenbaum, “Hierarchical deep reinforcement learning: Integrating temporal abstraction and intrinsic motivation,” 2016.

[6] S. Singh, A. G. Barto, and N. Chentanez, “Intrinsically motivated reinforcement learning,” 2005.

[7] R. Salakhutdinov and A. Mnih, “Bayesian probabilistic matrix factorization using markov chain Monte Carlo,” 2008, doi: 10.1145/1390156.1390267.

[8] T. Chen and W. Su, “Indirect Customer-to-Customer Energy Trading with Reinforcement Learning,” *IEEE Trans. Smart Grid*, 2019, doi: 10.1109/TSG.2018.2857449.

[9] C. A. Schroeder de Witt, J. N. Foerster, G. Farquhar, P. H. S. Torr, W. Böhmer, and S. Whiteson, “Multi-agent common knowledge reinforcement learning,” 2019.

[10] A. Raffin, S. Höfer, R. Jonschkowski, O. Brock, and F. Stulp, “Unsupervised Learning of State Representations for Multiple Tasks,” *Iclr*, 2017.

[11] Y. Bengio, A. Courville, and P. Vincent, “Representation learning: A review and new perspectives,” *IEEE Trans. Pattern Anal. Mach. Intell.*, 2013, doi: 10.1109/TPAMI.2013.50.

[12] O. Nachum, H. Lee, S. Gu, and S. Levine, “Data-efficient hierarchical reinforcement learning,” 2018.

[13] P. Henderson, W. Di Chang, P. L. Bacon, D. Meger, J. Pineau, and D. Precup, “Optiongan: Learning joint reward-policy options using generative adversarial inverse reinforcement learning,” 2018.

[14] A. Wilson, A. Fern, and P. Tadepalli, “Using trajectory data to improve Bayesian Optimization for Reinforcement Learning,” *J. Mach. Learn. Res.*, 2014.

[15] A. Karpathy and M. Van De Panne, “Curriculum learning for motor skills,” 2012.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

AE-DCNN: Autoencoder Enhanced Deep Convolutional Neural Network For Malware Classification

Shashank Kumar
Dept. of Computer Science & Engineering
Delhi Technological University
Delhi, India
shashank3256@gmail.com

Shivangi Meena
Dept. of Computer Science & Engineering
Delhi Technological University
Delhi, India
shivangimeena99@gmail.com

Savya Khosla
Dept. of Computer Science & Engineering
Delhi Technological University
Delhi, India
savyakhosla08@gmail.com

Anil Singh Parihar
Dept. of Computer Science & Engineering
Delhi Technological University
Delhi, India
parihar.anil@gmail.com

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498570

***Abstract*—Malware classification is a problem of great significance in the domain of information security. This is because the classification of malware into respective families helps in determining their intent, activity, and level of threat. In this paper, we propose a novel deep learning approach to malware classification. The proposed method converts malware executables into image-based representations. These images are then classified into different malware families using an autoencoder enhanced deep convolutional neural network (AE-DCNN). In particular, we propose a novel training mechanism wherein a DCNN classifier is trained with the help of an encoder. We conjecture that using an encoder in the proposed way provides the classifier with the extra information that is perhaps lost during the forward propagation, thereby leading to better results. The proposed approach eliminates the use of feature engineering, reverse engineering, disassembly, and other domain-specific techniques earlier used for malware classification. On the standard Malimg dataset, we achieve a 10-fold cross-validation accuracy of 99.38% and F1-score of 99.38%. Further, due to the texture-based analysis of malware files, the proposed technique is resilient to several obfuscation techniques.**

***Index Terms*—Autoencoder, Deep Convolutional Neural Network, Malware Classification, Pattern Recognition, Security**

I. Introduction

Malware is a type of software purposefully designed to cause harm to the user's system without their consent or knowledge. This harm typically involves manipulating stored information, distributing unauthorized access, and leaking private information. Over time, our dependencies on computer systems have tremendously increased. From individuals to organizations like hospitals, banks, etc., we all store, perform operations and do a lot more on computer systems. That is why ensuring a safe computer system for everyone is essential. However, the number of malicious attacks on systems has been increasing tremendously. The 2019 report by McAfee Labs reflects this growth by stating that ransomware attacks increased by over 100% [1]. Similarly, Kaspersky Security Network (KSN) reported an increase of more than 70% in the number of users who were victims of malicious attacks devised to collect user's data without their knowledge [2]. The most dominant reason behind this vast number is that malware authors make minute changes in an existing malware to make it appear as a new one and thus elude security systems such as anti-viruses installed in our computers.

Malware analysis investigates the suspected file's intention, way of working, and resulting harm. There are two ways in which this analysis is done, namely Static analysis and Dynamic analysis. The former works by a signature-based method and does not execute the suspicious file. The latter executes the file inside an isolated system within a regulated environment. Apart from being time-exhausting and resource-intensive, malware with minor obfuscation is not correctly analyzed.

Similar to malware analysis, the classification of malware into different malware families can help determine their intentions, functioning, and likely aftermath. Lately, many machine learning and deep learning techniques have been developed to detect and classify malware. Some of these techniques [3], [4] depend on extensive feature engineering to build a database of malware features. However, as the adversaries generate new malware every day, it becomes unreasonable to update the database with new malware features continuously. Likewise, traditional machine learning methods [5]–[7] perform well when the dataset is small and the target classes are fewer. However, as the dataset's size increases and the number of target classes grows, traditional machine

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:51:30 UTC from IEEE Xplore. Restrictions apply.

learning models begin to show their limitations. They take much time to train and lack the competence to learn the highly multidimensional features required for accurate classification. Hence, they become impractical and cost-ineffective for real-world setup. More recently, researchers are turning to deep learning for malware classification [8]–[11]. These techniques have resulted in significant improvements over the traditional approaches in terms of classification time and accuracy. They use neural networks to project malware representations into a high-dimensional feature space and then classify them into their respective families.

In this paper, we propose a novel deep learning model, AE-DCNN, for malware classification. The architecture classifies the malware executables by first converting them into their respective image-based representations. Then, an encoder is pre-trained using the malware images. The encodings generated by the encoder are then used to boost the performance of a deep convolutional neural network. Due to the malware's texture-based analysis, the proposed technique is resilient to obfuscation techniques like section encryption and packing.

Following are the main contributions of this paper:

- We propose a novel deep learning-based model for performing the task of malware classification. The proposed model is simple, robust and computationally efficient. Further, it does not require feature engineering and domain expertise.
- We propose a novel autoencoder architecture for generating vector encodings of malware executables. The proposed architecture is simple and generates robust and expressive encodings from malware images.
- We present an innovative training mechanism in which a deep convolutional classifier is trained leveraging the encodings generated by an encoder. We empirically demonstrate that the mechanism leads to a significant boost in the classifier's performance.
- We perform a comprehensive evaluation of the proposed model and demonstrate the significance of each of its components.

II. Methodology

In this section, we elaborate on the proposed framework and technique. As shown in Fig. 1., the model's architecture contains the following three components:

- An Encoder, $E(\cdot)$ - It generates a vector encoding of the malware images.
- A Feature Extractor, $F(\cdot)$ - It is a deep convolutional neural network that extracts high-dimensional feature representations from the images.
- A Classifier, $C(\cdot)$ - It predicts the family of the malware files.

A. Data Preprocessing

The dataset used in this paper is the Malimg Dataset [12]. It contains 9339 data points corresponding to 25 classes. These data points are image-based representations of malware executables, and the classes are their respective malware families. The images are generated by combining the bits of malware executables into 8-bit vectors and organizing these vectors into a two-dimensional array. This 2D array can then be rendered as a grayscale image. Fig. 2 elaborates this process of generating image-based representations from executable files, and Fig. 3 shows a sample of image-based representation for each malware family.

B. Encoder, $E(\cdot)$

The encoder module is used to generate compressed vector encodings of the malware images. For our use case, we train an autoencoder on the malware images and use its encoder part to generate compressed encodings of the images. In particular, we construct a CNN-based encoder $E(\cdot)$ and a CNN-based decoder $D(\cdot)$ with a bottleneck layer in between. $E(\cdot)$ learns to compress the image into a vector encoding that is expressive enough that $D(\cdot)$ can use it to reconstruct the input image. This enables us to use the autoencoder's encoder to extract the compressed and expressive encoding, e, of the malware images. Fig. 4 shows the architecture of the proposed autoencoder.

Each block in the encoder part of the autoencoder consists of a convolution layer with kernel size 3×3 followed by a Max Pooling layer with a stride of 2×2. Likewise, each block in the decoder part of the autoencoder consists of a convolution layer with kernel size 3×3 followed by an upsampling layer.

The autoencoder is pre-trained on images of size 512×512. Once the autoencoder is trained, we discard the decoder part and use the encoder in the AE-DCNN framework (Fig. 1) for generating compressed malware encodings. We use the mean-squared error loss to train the autoencoder. Equation (1) gives the value of the mean-squared loss for the kth data point, and (2) gives the value of loss for a batch of N images.

$$L^k_{mse} = \frac{1}{H \times W} \cdot \sum_{i=1}^{H} \sum_{j=1}^{W} (x^k_{ij} - d^k_{ij})^2 \tag{1}$$

$$L_{enc} = \frac{1}{N} \cdot \sum_{k=1}^{N} L^k_{mse} \tag{2}$$

Here, H and W represent the height and width of the image, respectively. x^k_{ij} represents the pixel in the ith row and j^{th} column of the kth image. Likewise, d^k_{ij} represents the pixel in the ith row and j^{th} column of the image generated by the decoder for the kth sample.

C. Feature Extractor, $F(\cdot)$

In our framework, a feature extractor, $F(\cdot)$, is used to extract high-dimensional feature representations from the malware images. We use the ResNet50 model pre-trained on the ImageNet dataset as the underlying architecture of the feature extractor. ResNet50 is a CNN-based architecture that deploys a series of residual blocks containing skip connections. These skip connections address the issue of diminishing accuracy with the increasing number of layers in the model. Thus, residual blocks have been effective when training deep models.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:51:30 UTC from IEEE Xplore. Restrictions apply.

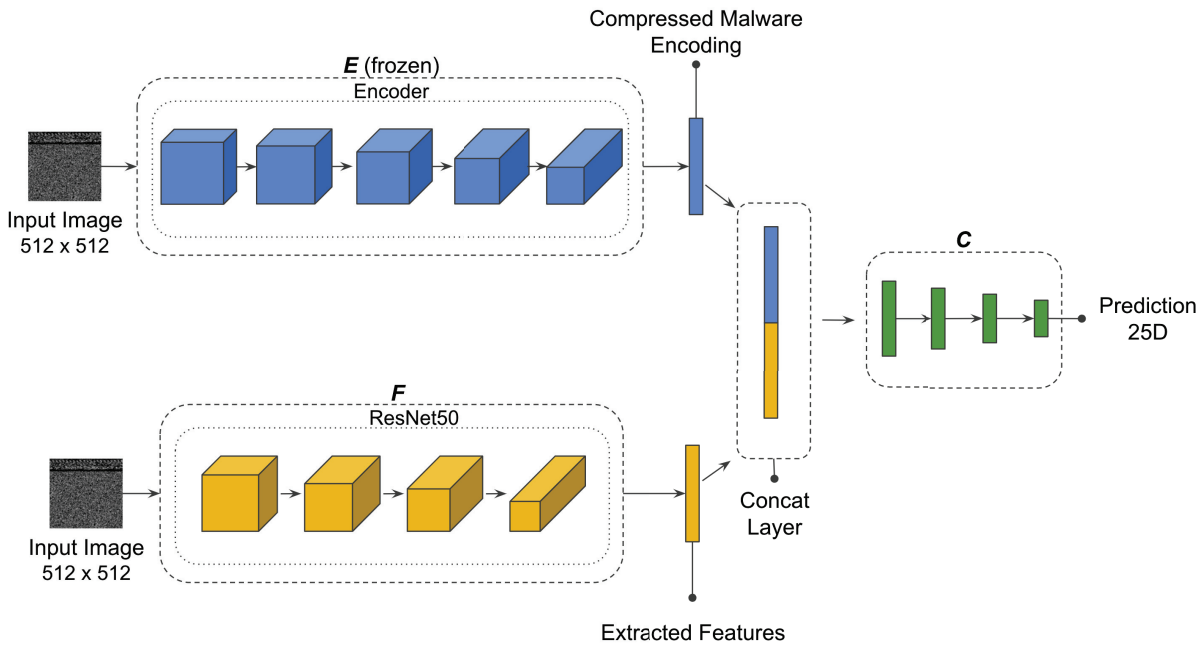


Fig. 1. The proposed AE-DCNN framework

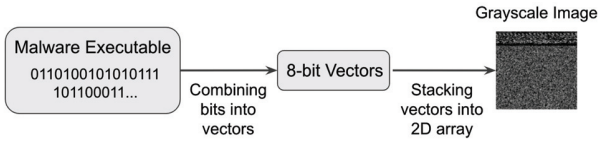


Fig. 2. Procedure for generating images from malware binaries

Also, ResNet50 has achieved high accuracy on many transfer learning tasks. These attributes make it an effective feature extractor.

D. Classifier, $C(\cdot)$

The classifier, $C(\cdot)$, is a multi-layered perceptron that takes the concatenated output of the encoder and the feature extractor as input and outputs the final predictions. We use a 4-layered fully-connected network with layers of size 512, 256, 128, and 25. The activation function for the hidden layers is ReLU, and for the output layer, it is SoftMax. The classifier is trained using the categorical cross-entropy loss given by (3).

$$L_{cce} = \sum_{i} y_i \cdot log(p_i) \tag{3}$$

Here, L_{cce} represents the categorical cross-entropy loss, y_i represents the true class label of ith data sample, and p_i denotes the model's prediction for i^{th} sample.

E. Training Algorithm

We introduce a novel training procedure wherein an encoder is used to enhance the training of the feature extractor and the classifier. Algorithm 1 gives a detailed overview of the entire training mechanism.

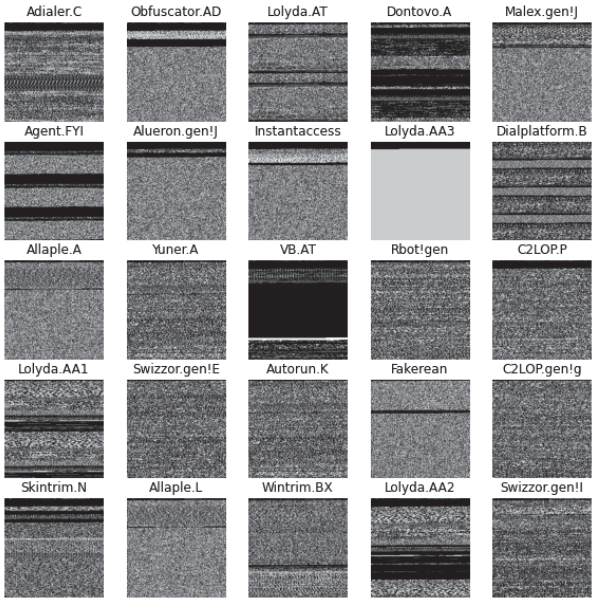


Fig. 3. Image-based representation of malware binaries

First, the autoencoder, $AE(\cdot)$, is independently trained on the malware images, x, using the L_{enc} loss defined in (1). Once trained, the decoder, $D(\cdot)$, is discarded, and the weights of the encoder, $E(\cdot)$, are frozen. In the algorithm, we use e and d to represent the outputs of the encoder and the decoder, respectively.

Second, the feature extractor, $F(\cdot)$, and the classifier, $C(\cdot)$, are trained end-to-end using the L_{cce} loss defined in (3). In

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:51:30 UTC from IEEE Xplore. Restrictions apply.

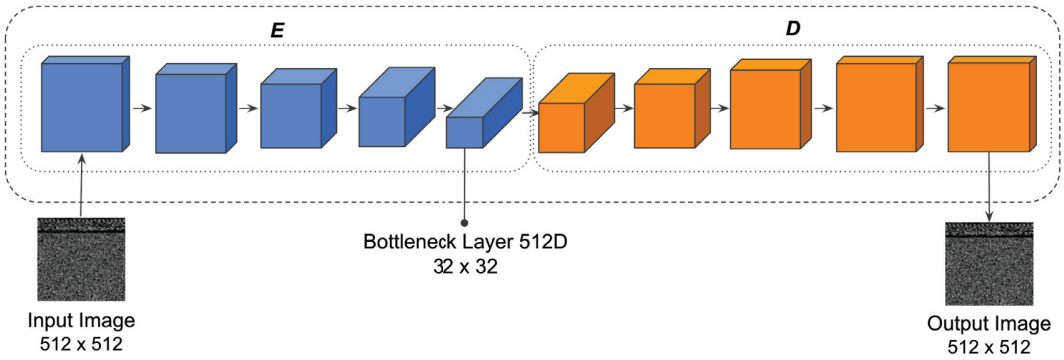


Fig. 4. The architecture of the proposed autoencoder

Algorithm 1 The algorithm for training AE-DCNN for malware classification

Require: The training set $DS_{train} = (x, y)$, Autoencoder $AE(\cdot)$ = [$E(\cdot)$, $D(\cdot)$], Feature Extractor $F(\cdot)$, Classifier $C(\cdot)$, number of epochs to train autoencoder $epoch_ed$, number of epochs to train the feature extractor and the classifier $epoch_fc$.

Initialize $AE(\cdot)$ = [$E(\cdot)$, $D(\cdot)$], $F(\cdot)$, and $C(\cdot)$

for $epoch$ = 1 to $epoch_ed$ **do**
  for all $(x, _)$ in DS_{train} **do**
    $e \leftarrow E(x)$;
    $d \leftarrow D(e)$;
    $l_enc \leftarrow L_{enc}(x, d)$;
    Update weights of $E(\cdot)$ and $D(\cdot)$ using l_enc
  end for
end for

Freeze the weights of $E(\cdot)$ and discard $D(\cdot)$.

for $epoch$ = 1 to $epoch_fc$ **do**
  for all (x, y) in DS_{train} **do**
    $e \leftarrow E(x)$;
    $f \leftarrow F(e)$;
    $i \leftarrow concat(e, f)$;
    $p \leftarrow C(i)$;
    $l_cce \leftarrow L_{cce}(y, p)$;
    Update weights of $F(\cdot)$ and $C(\cdot)$ using l_cce
  end for
end for

this step, the output of the feature extractor (f) and the output of the encoder (e) are concatenated to form the input (i) for the classifier. The classifier's output (p) forms the probability distribution of the predicted malware families.

III. Results and Discussions

We evaluate the performance of AE-DCNN on the Malimg dataset using accuracy and F1-score as the evaluation metrics. Equation (4) and (7) give the formula for the two, respectively.

$$Accuracy\ (A) = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$Precision\ (P) = \frac{TP}{TP + FP} \tag{5}$$

$$Recall\ (R) = \frac{TP}{TP + FN} \tag{6}$$

$$F1 - score\ (F) = \frac{2 \times P \times R}{P + R} \tag{7}$$

Here, TP, TN, FP, and FN are the number of true-positive, true-negative, false-positive, and false-negative predictions respectively.

Since the dataset is small in size and there is no pre-defined train-test split, we used the stratified 10-fold cross-validation to evaluate the performance of the model. For 10-fold cross-validation, we randomly divide the dataset into ten groups. Then the model was trained and evaluated ten times such that each group got to be the testing set once while the other nine combined to form the training set. Equation (8) gives the formula to compute the 10-fold cross-validation evaluation metric.

$$E = \frac{1}{10} \times \sum_{i=1}^{10} E_i \tag{8}$$

Here, E represents the evaluation metrics i.e., $E \in \{A, F\}$, and E_i is the value of the evaluation metric for i^{th} fold.

In order to establish the significance of the encoder module, we conduct two sets of experiments - one with the vanilla ResNet50 trained using standard training procedure and another with ResNet50 aided by autoencoder for its training. Table I summarises the results.

As can be seen from the table, ResNet50 trained with the aid of an autoencoder shows a significant increase in accuracy and F1-score when compared with the vanilla ResNet50 model. In particular, we observe a gain of 0.58% in accuracy and a gain of 0.64% in the F1-score. Since the baseline results are high, this is a significant gain. Further, due to the compact

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:51:30 UTC from IEEE Xplore. Restrictions apply.

TABLE I
PERFORMANCE OF AE-DCNN ON THE MALIMG DATASET

Model	Accuracy	F1-score
ResNet50	98.80%	98.74%
ResNet50 + Autoencoder	99.38%	99.38%

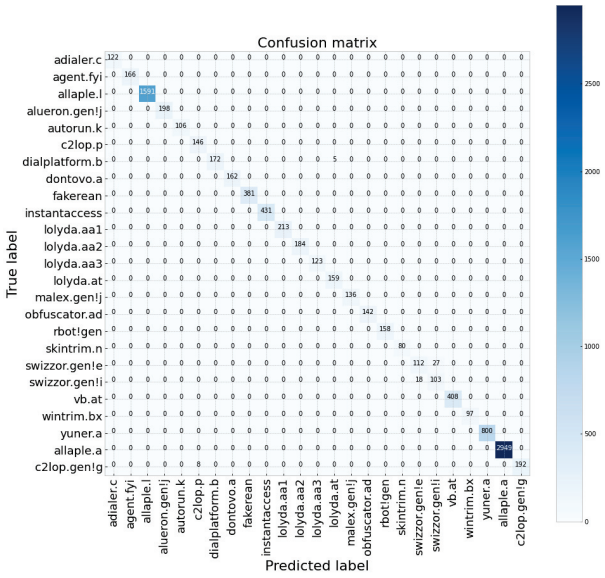


Fig. 5. Confusion Matrix for AE-DCNN

architecture of the autoencoder, the increase in parameter count is insignificant. We speculate that using autoencoder in the proposed way, we can provide the classifier with the extra information that is perhaps lost while forward propagating through the ResNet50 model.

Fig. 5 shows the performance of AE-DCNN using a confusion matrix. As can be seen, the model got a little confused between Swizzor.gen!E and Swizzor.gen!I. This is because of a high level of similarity between samples of these classes. Further, it misclassified some samples from Swizzor.gen!I class as C2lop.P and some samples of C2lop.gen!G class as C2lop.P. Other than these, the model performed well for all other classes. These results also demonstrate that AE-DCNN is resilient to obfuscation. For example, despite the multilayered encryption of the code section in Allaple.L and Allaple.A, the model is able to classify all samples of these classes correctly, thereby demonstrating that it can handle sectional encryption in malware files. Similarly, AE-DCNN is able to classify all samples of the obfuscator.ad class correctly.

All our experiments are conducted on NVIDIA Tesla T4 GPU.

IV. CONCLUSION

This paper presented a novel deep learning model, AE-DCNN, and an innovative training mechanism for malware classification. The proposed method first transforms the malware executables into their respective image-based representations. The encoder is pre-trained on these malware images, and its encodings are then used to enhance the performance of a deep convolutional classifier. Through experimentation, we demonstrate the effectiveness of the proposed technique. High 10-fold cross-validation results show that the model is robust and highly generalizable. Further, the results show that our approach is resilient to the basic obfuscation technique.

REFERENCES

[1] C. Beek, T. Dunton, J. Fokker, S. Grobman, T. Hux, T. Polzer, M. R. Lopez, T. Roccia, J. Saavedra-Morales, R. Samani, and R. Sherstobitof, Mcafee labs threats report, 2019 (accessed on April 6, 2021), URL: https://www.mcafee.com/enterprise/en-us/assets/reports/rp-quarterly-threats-aug-2019.pdf

[2] Kaspersky, Protecting your personal data online at every point, 2020 (accessed on April 6, 2021). URL: https://media.kasperskydaily.com/ wp-content/uploads/sites/92/2020/01/27103216/ International-Privacy-Day-2020-Kaspersky-report.pdf

[3] J. Zico Kolter and Marcus A. Maloof. 2006. Learning to detect and classify malicious executables in the wild," J. Mach. Learn. Res., vol. 7, pp. 2721–2744, January 2006.

[4] G. Yan, N. Brown, and D. Kong, "Exploring discriminatory features for automated malware classification," Detection of Intrusions and Malware, and Vulnerability Assessment, Springer Berlin Heidelberg, pp. 41–61, 2013.

[5] K. Rieck, T. Krueger, and A. Dewald, "Cujo: Efficient detection and prevention of drive-by-download attacks," Proceedings of the 26th Annual Computer Security Applications Conference, ACSAC'10, Association for Computing Machinery, pp. 31–39, New York, USA, 2010, doi: 10.1145/1920261.1920267.

[6] K. Rieck, P. Trinius, C. Willems, and T. Holz_aff2n3, "Automatic analysis of malware behavior using machine learning," J. Comput. Secur., pp. 639–668, vol. 19, December 2011.

[7] L. Nataraj, V. Yegneswaran, P. Porras, and J. Zhang, "A comparative assessment of malware classification using binary texture analysis and dynamic analysis," Proceedings of the 4th ACM Workshop on Security and Artificial Intelligence, AISec' 11, Association for Computing Machinery, pp. 21–30, New York, USA, 2011, doi: 10.1145/2046684.2046689.

[8] M. Kalash, M. Rochan, N. Mohammed, N. D. B. Bruce, Y. Wang, and F. Iqbal, "Malware classification with deep convolutional neural networks," in Proceedings of the 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS), pp. 1–5, 2018, doi: 10.1109/NTMS.2018.8328749.

[9] Z. Cui, F. Xue, X. Cai, Y. Cao, G. Wang, and J. Chen, "Detection of malicious code variants based on deep learning," IEEE Transactions on Industrial Informatics, vol. 14, no. 7, pp. 3187–3196, July 2018, doi: 10.1109/TII.2018.2822680.

[10] M. Nisa, J. H. Shah, S. Kanwal, M. Raza, M. A. Khan, R. Damaše29 Vičius, and T. Blažauskas, "Hybrid malware classification method using segmentation-based fractal texture analysis and deep convolution neural network features," Applied Sciences, p. 4966, vol. 10, no. 14, 2020, doi: 10.3390/app10144966.

[11] D. Vasan, M. Alazab, S. Wassan, B. Safaei, and Q. Zheng, "Image-based malware classification using ensemble of CNN architectures (IMCEC)," Computers & Security, p. 101748, vol. 92, 2020, doi:10.1016/j.cose.2020.101748.

[12] L. Nataraj, S. Karthikeyan, G. Jacob, and B. S. Manjunath, "Malware images: Visualization and automatic classification," Proceedings of the 8th International Symposium on Visualization for Cyber Security, VizSec' 11, Association for Computing Machinery, New York, USA, 2011, doi:10.1145/2016904.2016908.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:51:30 UTC from IEEE Xplore. Restrictions apply.

An Analysis on Abrupt Fall of Indian Stock Markets

Sadhvi Mehra
Department Of Applied Mathematics
Delhi Technological University
Delhi, India
sadhvimehra1998@gmail.com

Sumedha Seniaray
Department Of Applied Mathematics
Delhi Technological University
Delhi, India
sumedhaseniaray@dtu.ac.in

2021 International Conference on Information Technology (ICIT) | 978-1-6654-2870-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICIT52682.2021.9491650

***Abstract*—Stock Market analysis and prediction has always been a daily research article for investors, traders, economists. A lot of trend analysis and studies of forecasting has been performed already using time series and machine learning models which aim to determine the behaviour of the market for the upcoming days. In this project, we propose a model which very well explains the fall Indian Market witnessed during the last quarter of 2019- 20 at the time of sudden outbreak of a pandemic. We used NIFTY 50(Index of Indian Markets)and also took into account the S&P500 (Index of USA) and SSE(Index of China) to highlight the fact how the relation of global markets affected Indian Markets.**

***Index Terms*—market;index;logreturn;GrangerCausalitytest; VAR;RelativeOrdering;Seasonality;Stationarity;AIC.**

I. Introduction

Capital markets play an important role in a country's economic growth in the modern period. The stock market is divided into two categories: main and secondary markets. The primary market is where financial instruments are first made available for trading; the secondary market, also known as stock exchanges, deals with publicly traded securities that have already been sold in the primary market.Stock Exchanges are one of the most important aspects of the stock market and are used to gauge an economy's economic health. The stock market allows surplus funds to flow freely to industry in the form of capital, boosting economic growth. The NIFTY 50 Index is the benchmark stock market index for the Indian share market. The NIFTY 50 index covers 13 sectors of the Indian economy and provides investment managers with a single portfolio that gives them exposure to the Indian market. Forecast of stock pattern has for quite some time been a charming point,what's more, is widely concentrated by specialists from various fields. Time Series,a grounded calculation in a wide reach of uses, has been widely read for its possibilities in expectation of monetary business sectors and financial markets.

Nonetheless, in a significant number of these published works, the highlights chosen in machine learning algorithms are often forgotten among the knowledge contained within a similar market under consideration. This distinction leaves out important data transmitted by various substances, making forecast outcome less precise.Endeavors have been done to break the limits by joining outer data through new monetary news.

We propose using global stock data, along with information on other monetary objects, as input features to our algorithm in this project. In specific, we are keen on the connection between the closing costs of the market indices of India and the closing price of US markets as well as closing costs of Indian markets and that of China markets.Chinese items structure a basic piece of the production network for firms in numerous areas in India. Key areas influenced being: Pharmaceuticals: India gets 70% of dynamic drug fixings or APIs utilized in drugs from China and a disturbance in that would antagonistically influence the area. In 2020,due to the pandemic situations several online social Campaigns like boycott Chinese goods led to less demand of Chinese products. Our theory is supported by numerous results, and we believe,the information from overseas stock, especially one with a strong transient relationship with the Indian index, should be valuable.

The rest of the project is organised as follows: Section II delves further into the literature review. In Section III, numerical results are presented, accompanied by an analysis and discussion including the proposed algorithm's main concept, data collection, and exploratory data analysis.Section IV,sets up a basic machine learning model to show the capacity of the proposed solution . Section V sums up the entire report.

II. Related Works

A Comparative Study and Analysis of Time Series Forecasting Techniques by Srihari Athiyarath, Mousumi Paul, Srivatsa Krishnaswamy discusses how many practical domains have a lot of time series data.The study of such data is extremely useful in a variety of applications that witness changing patterns or particular requirements. Energy consumption, business product prices, and all other areas that exhibit seasonality. By comprehending the time series' essence and nature of time series the given paper described various methods to gain valuable information which meets business needs, depending on the essence of the time series and goal of study. This paper examines and contrasts different algorithmic forecasting methods, their weaknesses and utility for different kinds of time series data in various fields.

Stock Market Forecasting Using Machine Learning Algorithms by Shunrong Shen, Haomiao Jiang. The

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 02,2021 at 05:07:10 UTC from IEEE Xplore. Restrictions apply.

developers proposed a new forecasting algorithm based on the use of SVM that forecast next-day price trend by exploiting the relationship between global capital markets and different financial instruments over time. Finally, they built a straightforward trade model to compare the performance of the suggested prediction solution to other benchmarks.

Time Series: Economic Forecasting by M. W. Watson A variety of economic activities, including monetary and fiscal policy, use time-series forecasts. The most critical aspect of economic forecasting is choosing the appropriate prediction model(s) for the nature of the problem. X_t provides several time-series which could be used in a variety of ways lead to forecasting y_{t+1} in multivariate time-series models. Both scientific experience and economic theory typically affect the selection of these series. A difference between short and long - term interest rates, for example, is sometimes a helpful predictor of future inflation, according to interest rate term structure theory.Determining and communicating its degree of uncertainty in a prediction, as well as safeguarding against model instability.

Stock Market Movements and linkages between emerging markets in Asia and developed market indices: Short Run and Run Analysis.The share price fluctuations and relations between two developed markets and Asian developing economies (China, India, Indonesia, Korea, Malaysia, Philippines, Taiwan, and Thailand) are investigated in this research paper (i.e. USA and Japan). This study employs statistical analysis, the unit root test, association, and the pairwise Granger Causality test. The study examined both short-term (year by year) and long-term (over the entire study period) movements and linkages between Asian emerging stock markets and two existing stock markets using daily data for a period of almost ten years. There is a short-term relationship between these markets,a strong long-term relationship is absent.
Study's key Objectives: The following are the research study's goals.
1. Examining over short and long time periods, the normality and stationarity of sample indices from emerging Asian markets and developed markets.
2. Investigating over short and long time periods, co-movements and dynamic linkages between selected emerging stock market indices in Asia and selected developed markets.

Employing deep learning methods for predicting helpful reviews by Abdalraheem Alsmadi,Shadi AlZu'bi,Mahmoud Al-Ayyoub,Yaser Jararweh. Using three approaches: supervised, semi-supervised, and pre-trained models, the researchers used a series of machine/deep learning methods to classify helpful reviews of E-commerce websites (Amazon) in this analysis. The dataset included 83.7 million exclusive reviews from 24 major product categories. One of the topics addressed in this report was the distribution of input into six classes based on the number of votes received: 0 votes, 1 to 5 votes, 5 to 10 votes, 10 to 50 votes, 50 to 100 votes, and more than 100 votes. Only if a review receives more than 75% of the total votes is it considered beneficial. The semi-supervised model outperforms the supervised model in all categories. The findings show that working with pre-trained models like BERT and RoBERTa (Pre Trained Models) significantly improved accuracy as compared to previous results, with the BERT model outperforming all supervised classifiers and the RCNN supervised approach in particular. The paper addresses sentimental analysis and the importance of reviews. Since a twitter storm can engage a large number of people and has the ability to lift or lower a stock, crypto currency and stock market investors often use tweets as one of the approaches to decide market sentiment.

An Efficient Cold Start Solution for Recommender Systems Based on Machine Learning and User Interests. by Bilal Hawashin,Shadi Alzubi,Ala Mughaid,Farshad Fotouhi,Ahmad Abusukhon.When a user creates an account, he is prompted to provide some details such as age, gender, occupation, and so on. This paper proposed using machine learning to discover trends that connect user profile information to user vested interests. The qualified classifier is used to make a decision about a new user's preferences when he registers and fills out his own profile information. The binary relevance with J48, Bayesian classifier chain, binary relevance with random forest, and deep neural network with back propagation were the four classifiers used to make recommendations. The binary relevance approach was chosen as the most effective training classifier, and J48 was used as a classifier. In terms of planning, the Bayesian Classifier Chain was also a time saver. The analysis suggests how the relationship between the user's and his interests is used to make recommendations.

III. Methodology

A. Basic principles

Globalization strengthens ties between stock markets all over the world.The economic collapse ,caused by the recent pandemic has sent a shock wave around the world.Almost every country's economy has suffered, and all important stock records have been shattered.There isn't a single financial sector that isn't related these days. Financial details, political tensions, and other international undertakings may trigger big swings for domestic stock prices.

Hence,we propose to use world significant stock prices for our ML-based predictor in this project. Specifically, the oversea business sectors that shut right before or near the start of Indian market exchanging day can provide useful information on the trend of the upcoming Indian trading day, as their current developments reflect possible market assumptions on the most latest market news or a response to major world events.Trade model proposed works on the principle of VAR-a time series model.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 02,2021 at 05:07:10 UTC from IEEE Xplore. Restrictions apply.

Vector AutoRegression(VAR): The Vector Auto Regression (VAR) time series model is an advanced statistical model which aims to collect the development as well as linkages amongst US, Indian and China market indices which are constructed as three individual time series. Every one of the variable in the given model are treated equally by providing a condition clarifying each variable's development in relation to it's own lags as well as to lags of other variables in model by ACF and PACF.VAR technique for displaying offers various benefits over conventional choices dependent on primary system of equations.This is both increasingly conservative in its use of data and provides theoretical advantages against structural representation.

B. Data Collection

The project is based on the daily stock market data of 4 years from 2017 – 2020. The data for the project is taken from the yahoo Finance. The main task of the project is to fit a time series models to the closing price index of the Indian market along with China and USA, and to study the behaviour of the Opening Price based on the past data. We also try to study the dependence of the Indian market on the USA market and China Market.Since it is a stock market data the main difficulty lies in the fact that the data doesn't follow any such pattern or law and the randomness present in the model is high. We try to explain as much variability as possible. We use our knowledge of time series modelling and Machine Learning techniques and forecast the Indian market index for the next few days.

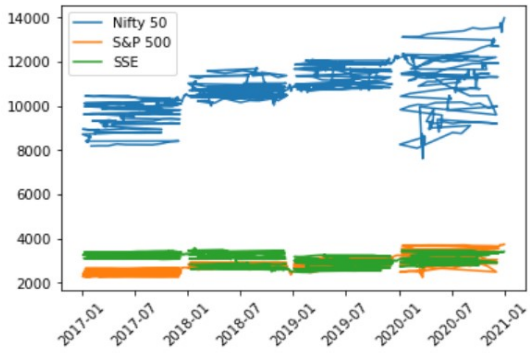


Fig1. Market Indices

The stock indices vary drastically.The Indian market index is seen to be extremely high ,however this is absurd .The reason is the exchange rate,before working on any model it is foremost to make the indices comparable. It covers daily price from 01-Jan-2017 to 31-Dec-2020 of all the three indices:We consider log returns on data as a basis for data consistency because markets remain suspended on holidays that differ by region, and missing values in other sources of data is dropped.Luckily our data contains only a few missing values so dropping null values seems a wise choice to enable continuity.

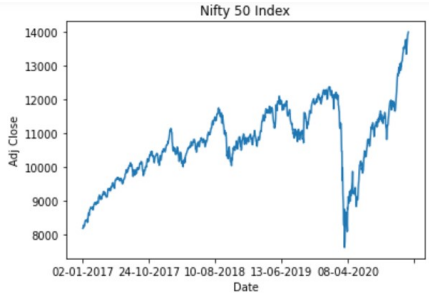


Fig2.Nifty 50

Before fitting of any time series model, we began our analysis by calculating descriptive of the data and visualizing the plot for presence of any systematic patterns. Table given summarizes the results obtained from this descriptive analysis of Nifty 50 data.

Count	983.00
Mean	10725.05
Std.Deviation	1081.42
Min	7610.25
1st Quartile	10071.75
Median	10768.05
3rd Quartile	11498.40
Max	13981.95

C. Trend and Randomness

In this project, we center around the analysis of trend (increasing or decreasing).As a result, the shift of a variable over time takes priority over the actual number of each variable. We detect the trend by Relative Ordering Test and try to remove it.Let the time series Nifty 50 be denoted by $\{X_1, X_2, ...X_n\}$
Define

$$q_{ij} = \begin{cases} 1, \text{ if } X_i > X_j \text{ when } i < j \\ 0, \text{ otherwise} \end{cases}$$

$$Q = \sum_{i=1} \sum_{j=2} q_{ij}$$

where Q counts the number of decreasing points in time series.
Null Hypothesis H_0:There is no trend in time series against the Alternate hypothesis H_1:There is trend in time series.
Under the null hypothesis $E(Q) = \frac{n(n-1)}{4}$

In order to detect the trend we perform the test on all the three market indices namely Nifty 50,S&P 500,SSE Composite Index.We report a p value around 0.0 for all the data sets and the number of discordants for the Nifty 50,S&P 500,SSE Composite turns out to be 104903,59223,257642 respectively.However the expected number of discordants are 243295.5,252757.5,235467.5 which is much greater than the actual number of discordants.Thereby discarding the null hypothesis and indicating an increasing trend.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 02,2021 at 05:07:10 UTC from IEEE Xplore. Restrictions apply.

For the detection of randomness we proceed with the Turning Point test. The test is based on the following hypothesis Null Hypothesis,H_0: Data set is Random Alternative Hypothesis, H_1: Data set is non random. After performing the test, we see that the p-value for the test comes out as 0.00 for all the indices. So, we reject our Null hypothesis i.e. we accept that the data set is non random.The number of turning points in the raw data set for Nifty 50,SSE,S&P500 are found to be 455,484,523 respectively.

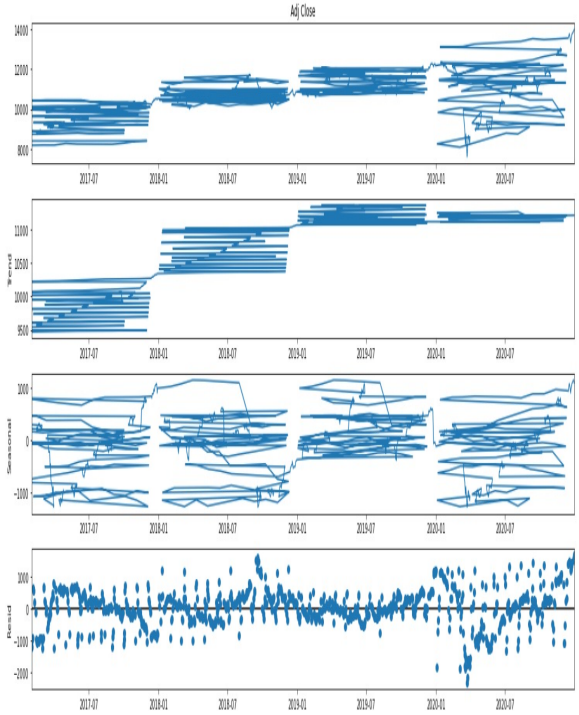


Fig3.Decomposition of deterministic Components of Nifty 50

A turning point is defined as peak when the value is greater than its 2 neighbouring values or a trough when a value is less than its 2 neighbouring values.All n-2 time points (2,3,..,n-1) are checked for being declared as turning points.Define

$$U_i = \begin{cases} 1,\text{if i is a turning point} \\ 0\ ,\text{otherwise} \end{cases}$$

P is the total number of turning points.i.e.$\sum_{n=2}^{n-1} U_i$
Null Hypothesis H_0 Series is purely random
Alternate Hypothesis H_1:Series is non-random.
Under the Null hypothesis

$$E(P) = \frac{2(n-2)}{3} \text{and } Var(p) = \frac{16n-29}{90}$$

Test Statistic :$Z = \frac{P-E(P)}{\sqrt{Var(P)}}$follow N(0,1) asymptotically.
We should reject null hypothesis of no trend at level of significance α if $|Z| > \tau_{\frac{\alpha}{2}}$,where $\tau_{\frac{\alpha}{2}}$is the $\alpha/2^{th}$ upper cut off point of standard normalisation.

Log Normalisation: Our first task was to omit null data points from the dataset and to bring it under same platform so that it is comparable. After we have successfully omitted the missing or 'NA' values from the dataset we are at a stage of merging the various market indexes and to study their dependence structure on each other. We assumed that the series is in a continuous time frame. We considered the log-return data for our entire analysis to make them comparable since every market index uses a different currency.

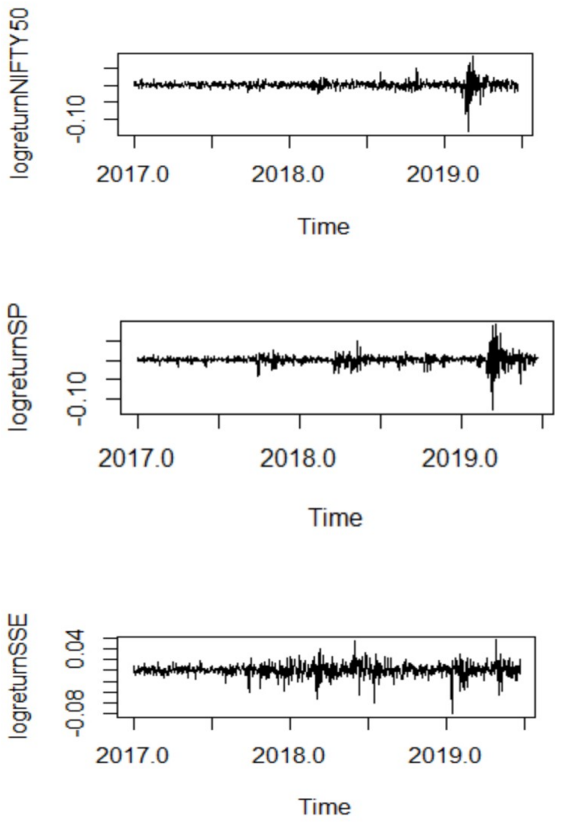


Fig4.Log Normalisation

Auto Correlation and Partial Auto Correlation

Now there exists a large number of models which can be fitted to any given time series data. However, a good way to proceed is to identify a small number of candidate models and then differentiate them using various model checking criterion. Towards that,we plotted the ACF and PACF plots for our detrended data. In both the plots we will look for the first lag for which the acf/pacf value lies outside the insignificant band. The first such lag from the ACF plot will be considered for the MA parameterwhile that from the PACF plot will be considered for the AR parameter. The first lag at which the value goes outside the insignificant band is equal to 5. Hence we will consider all the models with p and q both less than or equal to 5.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 02,2021 at 05:07:10 UTC from IEEE Xplore. Restrictions apply.

D. Multi Variate Analysis

The market indices tested now are detrended,random and stationary.More importantly they are comparable.Friedman test is used for detection of deterministic component i.e. seasonality.However the stock market is rarely found to be seasonal.

Correlation
Data sources like S&P 500 and SSE Composite Index stand as essential highlights for ML algorithm since they have a strong link to the Nifty 50. This understanding backs up our estimated guide, which we discussed in the previous section, about the interdependence of global markets and how the data reflected in their movements can be useful in forecasting Indian stock markets. This research paper has used two different ways to analyze the result first, the visual quality analysis, and the second is complexity analysis. Moreover, time complexity and space complexity parameters are measured to check the efficiency of the algorithm.We used **Granger Causality** to establish the inter dependence amongst the markets.Since the test works for bivariate analysis.We individually estimated the impact of either indices on Nifty 50 index and found out the desired results.

IV. Results and Discussions

Here as the aim of our project is to study the Indian market behaviour using India,China and USA stock Indices and then try forecasting the India open price using Both Indian,US market,and China market. So it is necessary to determine if USA stock market effect India and vice-versa. Thus, we have implemented two granger casualty test, one to check the effect of USA on India and the second one to determine if the India stock market has any effect on USA. Similiarly we check for China markets .We conduct both the tests at 5% significance and check the p-value for determining whether to accept or reject the null Hypothesis.
1.)H_0: log returns of USA do not cause log returns of India
H_1: not H_0
F-Test = 4.8566 , p-value = 0.0007053
At 5 % level of significance H_0 is rejected i.e. log-returns of USA do cause log-returns of USA.

2.) H_0: log returns of China do not cause log returns of India.
H_1: not H_0
F-Test = 2.7468 , p-value = 0.04192
At 5% level of significance H_0 is rejected i.e. log-returns of China do cause log-returns of India.
Interpretation:Hence we use multi variate study to forecast India stock market open price with the help of USA and China stock market.

At first, we have split the data set into train and test data. We have taken 90% of the data and we have tried fitting all the three main time series model (first component being India data,the second component being that of USA data and the third being that of China market) namely:
VAR(p)– Vector Auto regression Model of order p
VMA(q)- Vector Moving Average Model of order q

Fitting of VAR(p)

A VAR(p) model is given by:
$X_t = \mu + \psi_1 X_{t-1} + \psi_2 X_{t-2} + ... + \psi_p X_{y-p} + \epsilon_t$ Where,in case of multivariate data, $X_t = \begin{bmatrix} X_{t,1} \\ X_{t,2} \\ X_{t,3} \end{bmatrix}$ and ψ_i is a 3x3 matrix of coefficients. and $\epsilon_t \sim$ VWN (0,$\sum$)
Note that here our $X_{t,1}$=log returns of the India stock at time point t
$X_{t,2}$= log returns of the USA stock at time point t.
$X_{t,3}$= log returns of the China stock at time point t.

Fitting of VAR (1) We at first have fitted VAR (1) model and the model estimates are:

Constants	India	USA	China
μ	0.0003974501	0.0005416403	7.341956e-05

AR (1)-matrix

	India	USA	China
India	-0.0875	-0.0238	0.0854
USA	0.0156	-0.2700	-0.0136
China	-0.0295	-0.0227	0.0157

AIC: -26.52906

Fitting of VAR (2) We have now fitted VAR (2) model and the model estimates are:

Constants	India	USA	China
μ	0.0003742513	0.0004274671	6.305697e-05

AR (1)-matrix

	India	USA	China
India	-0.0824	-0.0238	0.08775
USA	0.0276	-0.2389	-0.00397
China	-0.0293	-0.0232	0.01588

AR (2)-matrix

	India	USA	China
India	0.0566	0.001542	0.00263
USA	0.1165	0.118701	0.02648
China	0.0199	-0.000863	0.02044

AIC: -26.53789

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 02,2021 at 05:07:10 UTC from IEEE Xplore. Restrictions apply.

We see that the AIC increase with the increase in order of the VAR(p).We have fitted VMA models of different orders but AIC of VAR remains lower.Thus, on the basis of AIC, we see that among all the models fitted on our multivariate data, VAR (1) fits the data best, i.e. it gives the lowest AIC among the whole set.The vector autoregression (VAR) model is one of the most common, scalable, and straightforward models for multivariate time series analysis. It's a normal progression from univariate autoregressive model to multivariate dynamic time series The VAR model has proved to be particularly useful in the following situation- describing the complex behaviour of economic and financial time series.

Findings of Review

Proposed Equation So,fitted VAR (1) model is:

$$X_t = 0.0003974501 - 0.0875X_{t-1} - 0.0238Y_{t-1} + 0.0854Z_{t-1}$$

Where, X component is of Nifty 50, Y component is of S&P 500 and Z component is of SSE Composite Index.Forecasting for multivariate VAR (1) Model We have fitted the model based on first 90% of the data and present the forecast values (out of sample predictions) on 2020.

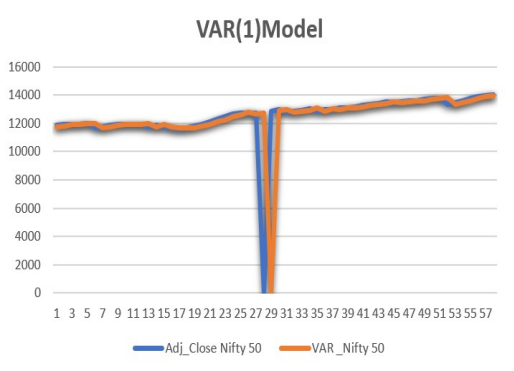


Fig5.Forecast of Indian Markets

The sharp dip in the figure 5 represents fall of the Indian markets and the revival beginning during July.

V. Conclusion

The study done for the fall of abrupt market that happened after the government of India announced lockdown during end march 2020 because of the Corona Pandemic was actually not solely based on India's decision and policy making.A major portion of trade is done by USA and China,These two countries with hold a big impact on Indian Economy.The two countries(China and India) traded about $92.68 billion worth of goods in 2019.

Since Corona already had spread widely in 2019 ending in major parts of China,the trade halted and stock markets there were already falling.Since we took into account the closing price of both the USA and China Indices,the markets indices were already running low during the end of 2019.This is the major reason although the Nifty 50 was high during the second last quarter of 2019 and beginning of 4th quarter of 2020 ,it began to decline even before the national lockdown that happened during the end of the quarter and the Indian economy suffered a setback in starting of financial year 2020-2021.

References

[1] M. W. Watson Time Series: Economic Forecasting.
[2] W. Huang et al., "Forecasting stock market movement direction with support vector machine," Computers Operations Research, 32, pp. 2513–2522005, 2005
[3] J. Moody, et al., "Learning to trade via direct reinforcement," IEEE Transactions on Neural Networks, vol. 12, no. 4, Jul. 2001.
[4] Srihari Athiyarath,Mousumi Paul, Srivatsa Krishnaswamy(2020).A Comparative Study and Analysis of Time Series Forecasting Techniques.
[5] S. Zemke, "On developing a financial prediction system: Pitfall and possibilities," Proceedings of DMLL-2002 Workshop, ICML, Sydney,Australia, 2002.
[6] Vatsal H. Shah, "Machine learning techniques for stock prediction,"
[7] Chavannavar, M. B., Patil, S. C., Simoes, M. (2016). Monetary Policy Effect on Nifty 50 and Sectoral Indices- A Study from Indian Stock Markets. International Journal of Latest Technology in Engineering, Management Applied Science, 59-69.
[8] Chowdhury, P. R., A, A. (2018). Impact of exchange rate fluctuation on stock market volatility - a study to predict the economic scenario in India. International Journal of Pure and Applied Mathematics, 4309-4316.
[9] D, R., Raju, J. K., G, B. K. (2016). An Impact of Currency Fluctuations on Indian. International Journal of Application or Innovation in Engineering Management, 146-151.
[10] Lioudis, N. K. (2018, April 9). Retrieved September 12, 2018.
[11] Patel, D., Kagalwala, N. (2013). The Impact of Exchange Rate on Indian. INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH, 1-2.
[12] Sekhri V, Haque M. (2015). Impact of Foreign Investments on Indian Stock Market. Asian Journal of Research in Banking and Finance, 168-185.
[13] Employing deep learning methods for predicting helpful reviews. In 2020 11th ICICS (pp. 007-012).IEEE
[14] An Efficient Cold Start Solution for Recommender Systems Based on Machine Learning and User Interests. In 2020 SDS (pp. 220-225). IEEE.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 02,2021 at 05:07:10 UTC from IEEE Xplore. Restrictions apply.

International Journal of Environmental Science and Technology
https://doi.org/10.1007/s13762-021-03615-5

ORIGINAL PAPER

An efficient adsorbent for the removal of Zn^{2+} Cd^{2+} and Hg^{2+} from the real industrial effluents

D. Chauhan[1] · A. Kumar[1] · S. G. Warkar[1]

Received: 22 March 2021 / Revised: 23 June 2021 / Accepted: 15 August 2021
© Islamic Azad University (IAU) 2021

Abstract
Meta-benziporphodimethene (*meta*-BPDM) ligand immobilized in polyacrylamide/ carboxymethyl guar gum (PAM/CMG) hydrogel was used for selective removal of Zn^{2+}, Cd^{2+} and Hg^{2+} ions from water. This hydrogel appears to be a competitive sorbent and has been successfully applied to efficient removal of Zn^{2+}, Cd^{2+} and Hg^{2+} ions from water in the presence of others. Batch adsorption experiments with respect to pH, initial metal ion concentration and contact time were also investigated. The selectivity towards Zn^{2+}, Cd^{2+} and Hg^{2+} is due to *meta*-BPDM ligands and other functional groups incorporated in hydrogel. Adsorption on hydrogel was found to be pH dependent and maximum adsorption was found to be at pH 6.5. The sorption equilibrium for Zn^{2+}, Cd^{2+} and Hg^{2+} ions was achieved in 8 h, 12 h and 6 h contact time at room temperature, 250 rpm and 500 ppm initial metal ion concentration, respectively. The percentage removal of Zn^{2+}, Cd^{2+} and Hg^{2+} was found to be 78.8%, 67.6%, and 80.4%, respectively. Maximum removal was observed with respect to Hg^{2+} ($Hg^{2+} > Zn^{2+} > Cd^{2+}$) and this can be explained by the hard-soft acid base principle. Scanning electron microscopy, Fourier transform infrared spectroscopy, and Thermogravimetry analysis were employed to monitor changes that have occurred in the hydrogels after the sorption of metal ions. Inductively coupled plasma mass spectroscopy (ICP-MS) and atomic absorption spectroscopy (AAS) analysis were employed to estimate the residual metal ion concentrations in water. Parallel adsorption and comparative studies indicated that the hydrogel has better adsorption capacity and selectivity.

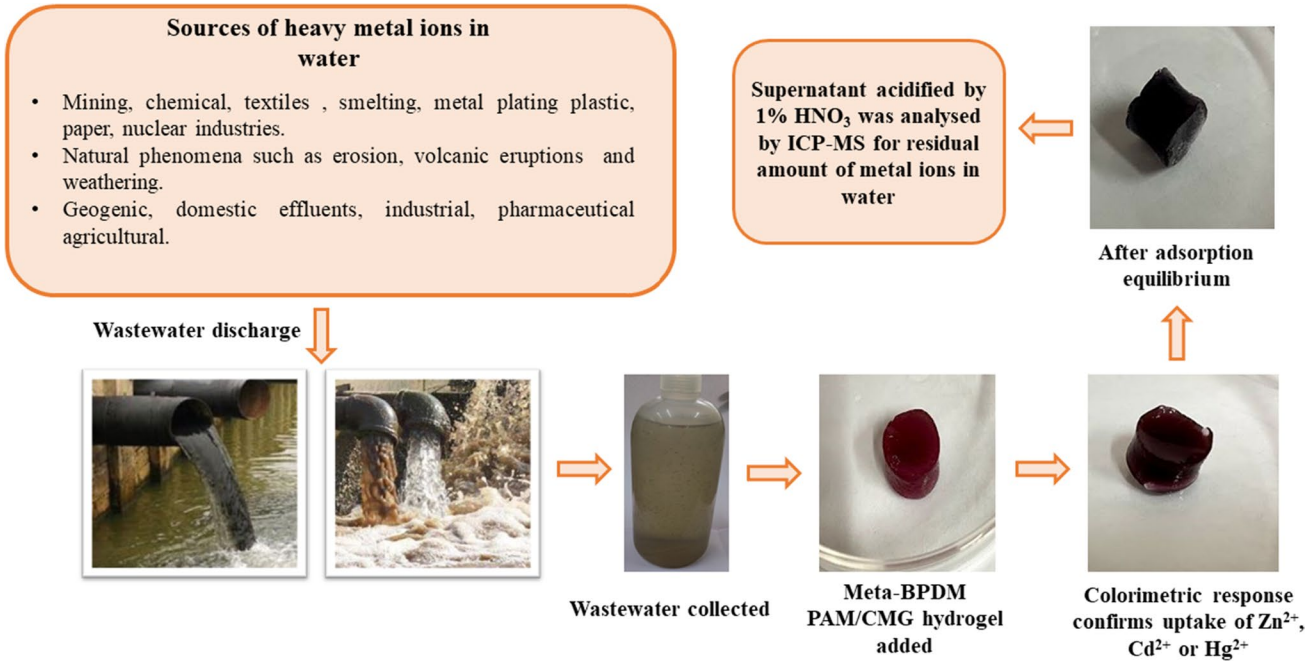
Editorial responsibility: Samareh Mirkia.

✉ A. Kumar
anil_kumar@dce.ac.in

[1] Department of Applied Chemistry, Delhi Technological University, Bawana Road, Delhi-42, India

Published online: 24 August 2021

Graphic abstract



Keywords Adsorption kinetics · Colorimetric sensor · Effluent · Hydrogel · Sorption

Introduction

Among the most common pollutants found in wastewater heavy metals are the main pollutants and causes severe ecological problems, thus their elimination from water during water treatment is crucial and requires research attention. The major pathway of heavy metal exposure to humans is food chain contamination. The main source of toxic heavy metal ions released in the environment are human and natural activities (Ahmad and Goni 2010). Unregulated discharge of non-biodegradable heavy metals from industrial, municipal wastewater and other sources brings significant amounts of toxic heavy metal ions. The wastewater contains toxic heavy metals released directly into aquatic ecosystems including rivers with or without pre-treatment (Lone et al. 2019). These metal ions may accumulate to very high toxic levels without any visible sign. Long-term irrigation with such heavy metal ion contaminated wastewater creates a problem for safe rational use of agricultural soils. Excessive agglomeration of heavy metals like Cd, Cu, Cr, Zn, Hg, Pb, Mn, etc., in agricultural soil affects food quality adversely. Plants and vegetables cultivated in the nearby zones of industrial areas show increased levels of heavy metals. This accumulation of heavy metal ions in edible and nonedible items is enough to cause clinical problems to both human beings and animals consuming that. Therefore, early detection, determination and elimination of heavy metal ions from effluents are important for the survival of living organisms (Awual 2015; Gupta et al. 2009).

Heavy metals are not biodegradable therefore their effective removal from aqueous medium is very important for the protection of public health and environment. Efficient techniques are required to monitor certain metal ions from water, mainly the toxic which are harmful even at low concentration (Jin and Bai 2002; Karbarz et al. 2018). Heavy metal ions are usually removed by physically and chemically from the wastewater via conventional process include coagulation/ flocculation, reverse osmosis, complexation/ sequestration, electrochemical operation, ion exchange, biological treatment, chemical precipitation, adsorption, membrane filtration, zeolites, carbon-based sorbent. The above-mentioned process becomes economically non-viable for the removal of heavy metals at lower concentrations and may cause secondary contamination (Meena et al. 2005; Namasivayam and Ranganathan 1995; Wang et al. 2013). Several techniques have been developed to eliminate heavy metal ions from wastewater but each one exhibits limitations for instance in case of solvent extraction the eluent has to be processed to avoid any further pollution. Chemical precipitation of toxic sediments is not an economically successful method, then regarding adsorption of pollutants on solid sorbents appeared to be an efficient technique. Among all this

adsorption technique is widely regarded as an easy, effective and versatile method for the removal of heavy metals from aqueous solution. Hydrogels are cross-linked polymers and have the ability of loading aqueous solutions and certain ions. The high-water retention capacity of hydrogel allows more rapid ion penetration into the cross-linked framework of hydrogel (Karbarz et al. 2018). They undergo reversible change in volume towards external stimuli such as change in pH and temperature. This property brought about the utilization of hydrogels as molecular detectors, absorbent, and sensors. The highly porous structure of hydrogel allows water molecules, ions to easily diffuse and interact with functional groups (-COOH, -NH_2, -OH) in a polymeric network. The potential of hydrogels to be complex with metal ions allows them to be an efficient metal absorbent (Lone et al. 2019).

Zinc is commonly used in the process of galvanization, alloys, industries, catalyst, textiles. The toxicity due to zinc is uncommon but excess ingestion may cause gastrointestinal distress and diarrhoea. Cadmium is commonly used in batteries, coating, mining alloy, smelting and plastic industry. It causes irritation to the respiratory tract and anaemia. Mercury is one of most toxic heavy metals and non-biodegradable in nature. Major sources of mercury in the environment are such as paint, pulp, fertilizer, thermometers, fluorescent light bulbs etc. It causes genetic defects, nervous system damage, impairment of pulmonary functions, mental retardation, and chest pain (Meena et al. 2005). Cd and Hg ions are toxic even at low levels and may lead to death after being accumulated at high concentration in the body. These are toxic, considered carcinogenic. There is great big demand for removal of such toxic heavy metal ions from wastewater (Karbarz et al. 2018).

Therefore easy, reliable, selective, and sensitive methods for Zn^{2+}, Cd^{2+} and Hg^{2+} ion monitoring and removal are important for environmental safety and health. Various other methods have been reported to remove toxic metal ions from water for instance, Cerium phosphate polypyrrole flower like nanocomposite material ($CePO_4$-PPY) was found to be a promising adsorbent for the efficient removal of Cr (VI) containing wastewater. Where removal process involved both adsorption and simultaneous in situ chemical reduction of Cr (VI) and to Cr (III) (Sahu et al. 2021a, b). Nanocomposite based on Lanthanum phosphate polyaniline ($LaPO_4$-PANI) was found to be promising adsorbent for the removal of Cr (VI). The adsorption mechanism involved the combined effect of both adsorption and reduction (Sahu et al. 2021a, b). The core–shell nanocomposite based on polyaniline (PANI) and thorium dioxide (ThO_2) used for the removal of Cr (VI) from water. ThO_2-PANI nanocomposite could effectively adsorb Cr (VI) and partially reduced to Cr (III) during the adsorption experiment (Sahu et al. 2019). Adsorbent based on lanthanum phosphate (LaP) and poly o-toluidine (POT) was found to be effective adsorbent with higher regeneration capability and potential removal of fluoride from ground contaminated water (Sahu et al. 2020). The detailed survey with few reported hydrogels-based adsorbent with removal efficiency are well presented in Table 1.

Methods for the removal of Zn^{2+}, Cd^{2+} and Hg^{2+} from water have been reported. For instance, DNA-functionalized polyacrylamide hydrogels employed for Hg^{2+}removal

Table 1 Showing Adsorption capacity and percentage removal of various heavy metal ions by hydrogel based adsorbents

S. no.	Hydrogel based adsorbents	Adsorption capacity q_e (mg/g)	% Removal	References
1	Spherical chitosan/gelatin hydrogel particles	Hg^{2+}-47.5, Pb^{2+}-7.62 Cd^{2+}-0, Cr^{3+}-1.5	Hg^{2+}-84.7, Pb^{2+}-8.7, Cd^{2+}-0 Cr^{3+}-6.7	Perumal et al. (2019)
2	Spherical shaped graphene oxide incorporated chitosan–gelatin hydrogel particles	Hg^{2+}-54.6, Pb^{2+}-5.4, Cd^{2+}-1.67, Cr^{3+}-0	Hg^{2+}-54.6, Pb^{2+}-7.3, Cd^{2+}-1.9, Cr^{3+}-0	Perumal et al. (2020)
3	Poly (acrylamide-co-sodium methacrylate) based hydrogel	Cu^{2+}-24.05 Cd^{2+}-33.0	Cu^{2+}-48, Cd^{2+}-66	Milosavljević et al. (2013)
4	Polyacrylamide based hydrogel	Cd^{2+}-5.3 mmol g^{-1}, Pb^{2+}-0.63 mmol g^{-1}, Zn^{2+}-1.27 mmol g^{-1}		Sharaf et al. (2009)
5	Polyacrylic acid hydrogel	Cd^{2+}-132.9 Cr^{6+}-58.1 Fe^{3+}-12.4 Mn^{2+}-120.4 Ni^{2+}-128.8 Ag^{+} and Ce^{3+}-203.5 Zn^{2+}-157.8	Cd^{2+}-57.1, Cr^{6+}-26.9, Fe^{3+}-5.3, Mn^{2+}-52.7, Ni^{2+}-52.5, Ag^{+}-45.3, Ce^{3+}-70, Zn^{2+}-58.4	Lv et al. (2019)
6	Gelatin/chitosan hydrogel		Hg^{2+}-97, Pb^{2+}-12, Cd^{2+}-2 Cr^{3+}-24	Lone et al. (2019)
7	Sulfonic acid based hydrogels	Cd^{2+}-0.95 Cu^{2+}-0.87 Fe^{3+}-0.83 Zn^{2+}-1.00 Mn^{2+}-0.77 Pb^{2+}-0.18		Atta et al. (2011)
8	Alginate-fibroid hydrogel	Cu^{2+}-316.0 Cd^{2+}-232.35 Pb^{2+}-465.2		Kong et al. (2020)
9	Graphene oxide/alginate based hydrogel membrane	Pb^{2+}-327.9 Cr^{3+}-118.6		Bai et al. (2020)

from water (Dave et al. 2010). A modified mesoporous conjugate adsorbent was formulated by fixing the organic ligand of 2-hydroxyacetophenone-4 N-pyrrolidine thiosemicarbazones onto mesoporous ZSM-5 for selective and effective mercury removal from aqueous solution (Abbas et al. 2018). Highly ordered mesoporous silica impregnated with ligand N,N-disalicylidene-4,5-dimethyl-phenylenedene-based optical nanocomposite material used for the sensitive and selective removal of $Hg^{2+}$from waste water (Awual 2017). Mesoporous silver-melamine-based nanocomposites formulated by controlled supramolecular self-assembly employed for selective removal Hg^{2+} (Liu et al. 2020). Poly(N-isopropylacrylamide) hydrogel cross-linked with the diacryloyl derivative of cysteine used for removal of Cd^{2+} and Pb^{2+} (Karbarz et al. 2018). A sensitive fluorescent sensor, i.e. 2,2-dipicolylamine modified naphthalimide fluorophore followed by immobilization to the surface of silica microsphere used for detection and removal of Cd^{2+}, Hg^{2+} and Pb^{2+} toxic metal ions from aqueous samples (Ding et al. 2015). Red mud, waste of an aluminium industry has been transformed into an efficient and inexpensive adsorbent for Zn^{2+}, Cd^{2+} (Gupta and Sharma 2002). Carbon aerogel is a new form of activated carbon used for adsorption Pb^{2+}, Cd^{2+} and Hg^{2+} metal ions (Kadirvelu et al. 2008). Rice straw was used as a biosorbent for adsorption of Cu^{2+}, Zn^{2+}, Cd^{2+} and Hg^{2+} ions in industrial effluents (Rocha et al. 2009).

This report presents a robust and cost-effective approach for removal of Zn^{2+}, Cd^{2+} and Hg^{2+} from aqueous solution using *meta*-BPDM embedded PAM/CMG hydrogel in pellet form. The sorption, adsorption kinetics and isotherm studies of this hydrogel has been reported here which was not mentioned so far. The presence of polar hydrophilic functional groups (-OH, NH_2, -COOH) in polyacrylamide/carboxymethyl guar gum and porphyrin analogue *meta*-BPDM act as binding sites via polar or electrostatic interaction. It has made *meta*-BPDM embedded PAM/CMG hydrogel an effective sorbent for Zn^{2+}, Cd^{2+} and Hg^{2+} in wastewater treatment. This study was carried out at Delhi Technological University (Delhi, India). All experiments were performed between 2020 and 2021.

Materials and methods

Chemicals required

Pyrrole, 2,6-Difluorobenzaldehyde and α, α dihydroxy-1, 3-diisopropylbenzene were purchased from TCI chemicals (India) Pvt. Ltd. Silica gel, Acetone, N, Ń-methylenebisacrylamide (MBA) and metal salts were obtained from Merck Germany. Boron trifluoride ethyl ether complex, 2,3-dichloro-5,6-dicyano-p-benzoquinone (DDQ) were obtained from SRL Pvt. Ltd. (India). Carboxymethyl guar gum was procured from Hindustan Gum and Chemicals Ltd., Bhiwani, Haryana (India). Potassium persulfate ($K_2S_2O_8$) was purchased from Thomas Baker chemicals and Acrylamide from SRL. Deionized water was used for synthesis and for all metal ions aqueous solution preparation.

Synthesis

meta-BPDM embedded polyacrylamide/carboxymethyl guar gum and polyacrylamide/carboxymethyl guar gum hydrogel were synthesized according to previous reports (Chauhan et al. 2020). The digital photo of hydrogel is shown in Fig. 1.

Instrumentation

The synthesized hydrogel and metalated hydrogel were characterized using Fourier-transform infrared spectroscopy (FTIR), thermogravimetric analysis (TGA), scanning electron microscopy (SEM). IR-spectrum was collected using KBr pellets containing 1% sample from 4000 cm^{-1} to 400 cm^{-1} (NICOLET 380), the surface morphology of hydrogel was observed by using SEM (JEOL JSM-6610LV) after coated with conductive film of platinum. The thermal stability of hydrogels was investigated by TGA (Perkin

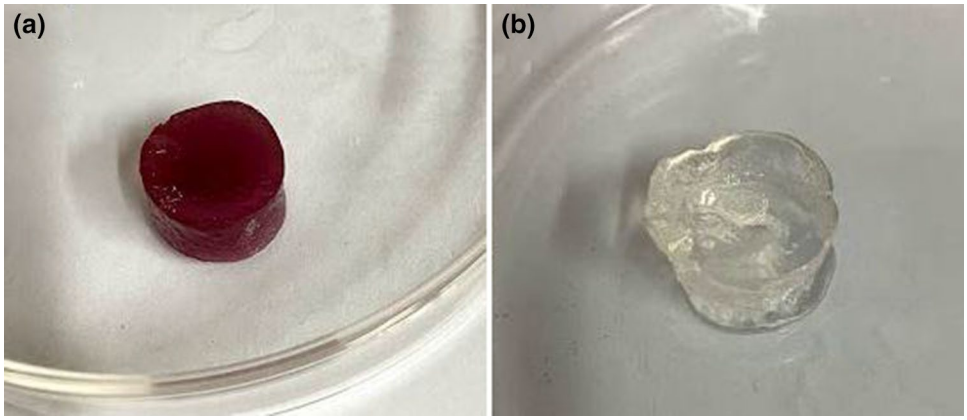


Fig. 1 **a** *meta*-BPDM PAM/CMG hydrogel, **b** PAM/CMG hydrogel

Elmer TGA 4000.). The sample was heated at 10 °C/min from 80 to 800 °C under the N2 atmosphere. The initial and residual concentration of metal ions in water were measured using ICP-MS (Perkin Elmer 300X) and AAS (ZX press 8000).

Selectivity studies

In a 50 mL solution of all metal ions with pH 6.5 and initial concentration 500 ppm, 0.082 g hydrogel pellet was added. The solution was shaken with 250 rpm at room temperature for 5 h. The equilibrium concentration of all metal ions was observed using AAS. The adsorption selectivity was also compared between Blank hydrogel, i.e. PAM/CMG hydrogel and *meta*-BPDM embedded PAM/CMG hydrogel. Removal percentage was calculated for each metal ion.

Adsorption experiments

Batch sorption of Zn^{2+}, Cd^{2+} and Hg^{2+} by hydrogel (0.082 g) were studied at room temperature (about 28 °C) using 10 mL of heavy metal ion aqueous solution. Analytical grade chloride salts of Zn^{2+}, Cd^{2+} and Hg^{2+} were used to prepare standard stock solutions in deionized water. Adsorption experiment was performed using a diluted stock solution at room temperature. The desired pH was adjusted using HCl (0.1 N) and NaOH (0.1 N). The effect of initial metal ion concentration, pH and contact time were studied. The results obtained from these studies were used to demonstrate the optimum conditions for maximum removal from aqueous solutions.

Removal of Zn^{2+}, Cd^{2+} and Hg^{2+} from industrial wastewater

0.082 g hydrogel pellet was suspended in 10 mL of industrial wastewater (collected from industries nearby DTU, Delhi). The pH of the mixture was adjusted to 6.5 and stirred for 4 h (250 rpm) at room temperature to obtain equilibrium. The initial and remaining concentration after adsorption of Zn^{2+}, Cd^{2+} and Hg^{2+} in the sample were determined by AAS.

Results and discussion

Synthesis and Characterization of blank, *meta*-BPDM embedded and metalated *meta*-BPDM embedded polyacrylamide/ carboxymethyl hydrogel

Fourier-transform infrared spectroscopy (FTIR) spectrums (Fig. 2) of *meta*-BPDM embedded polyacrylamide/ carboxymethyl hydrogel, Zn^{2+}-*meta*-BPDM PAM/CMG, Cd^{2+}-*meta*-BPDM PAM/CMG, Hg^{2+}-*meta*-BPDM PAM/CMG, Zn^{2+} Cd^{2+} and Hg^{2+}-*meta*-BPDM PAM/CMG. The absorption bands at 3200-3300 cm^{-1} were related to -OH and N–H stretching and N–H and O–H overlapped. The bands at 2730 cm^{-1} are due to saturated C-H and 2800 cm^{-1} unsaturated C-H stretching mode. The peak at 1050 cm^{-1} corresponds to C-N stretching, 1250 cm^{-1} represents C-O bending, 1450 cm^{-1} related to O–H bending, 1400 cm^{-1} corresponds to C–O–C stretching vibration in the CMG. The strong peak at 1720 cm^{-1} corresponded to C=O groups (Zhao et al. 2019).

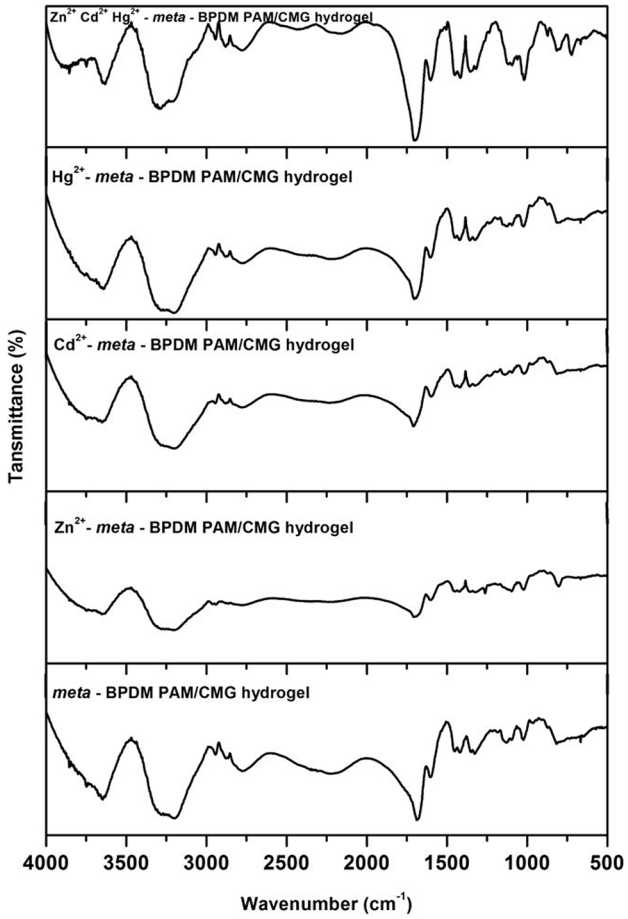


Fig. 2 FTIR Spectra of hydrogels

Thermogravimetric analysis was used to investigate thermal properties of hydrogel. Similar degradation pattern was observed with all hydrogels whereas in case of Zn^{2+} and Cd^{2+} metalated hydrogel decomposition is slightly different that may be due to binding coefficient of these two metal ions with *meta*-BPDM. Overall, it has been observed that insertion of *meta*-BPDM and its metalation do not affect thermal stability of the PAM/CMG hydrogel considerably, which indicates that there is no chemical interaction between them. The weight loss up to 100 °C was owing to water evaporation from the hydrogel. The hydrogel undergoes

sharp weight loss in the temperature range of 250–350 °C due to the degradation of -OH and carboxymethyl group. The weight loss from 350 to 500 °C is due to decomposition of the polymer backbone, crosslinking, and amide groups. The final weight loss is due to degradation of the organic backbone, C–O–C *meta*-BPDM ring. The overall analysis depicts that the hydrogel sample had adequate thermal stability and shown in Fig. 3.

Surface properties were analysed using scanning electron microscopy; the morphology was found to be irregular, porous which provides effective adsorption for the heavy metals. The morphology is presented in Fig. 4 (inset A), present surface microphotograph of *meta*-BPDM PAM/CMG hydrogel and shown in Fig. 4 (inset B), showing of

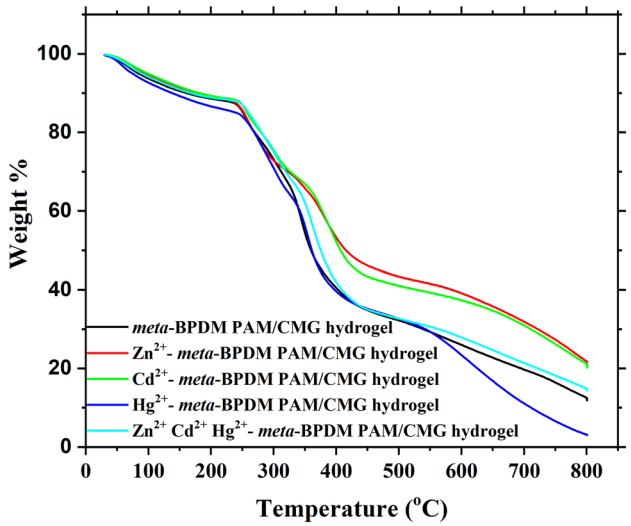


Fig. 3 TGA graphs of hydrogels

Zn^{2+}, Cd^{2+} and Hg^{2+} metal-loaded *meta*-BPDM PAM/CMG hydrogel. The surface morphology slightly changes after adsorption that may be attributed to conformation change on metalation. Overall, no chemical change in morphology is observed on metalation which indicates that there is physical interaction between metal loaded *meta*-BPDM and hydrogel.

Selectivity/competitive studies

The industrial discharge usually contains multiple metal ions therefore it has become important to study the competitive removal of metal ions. Selective removal of Zn^{2+}, Cd^{2+} and Hg^{2+} was observed in a mixture of multiple metal ions and shown in Fig. 5. It can be attributed to specific interaction of Zn^{2+}, Cd^{2+} and Hg^{2+} with embedded *meta*-BPDM or active functional groups in the hydrogel. Other than *meta*-BPDM a large number of $-NH_2$, -OH, -COOH, and $-CONH_2$ functional groups of PAM/CMG hydrogel can act as binding sites for the Zn^{2+}, Cd^{2+} and Hg^{2+}. Functional groups having nitrogen are selectively involved in Hg^{2+} binding. Therefore, adsorption isotherm and kinetics was studied with respect to only Zn^{2+}, Cd^{2+} and Hg^{2+} removal. The aqueous solutions of all metal ions of required concentrations were prepared using their chloride salts in deionized water. The higher affinity of hydrogel among Zn^{2+}, Cd^{2+} and Hg^{2+} was found to be with respect to Hg^{2+}. This can be supported by hard soft acid base theory towards active functional groups having nitrogen and *meta*-BPDM. The reason for higher chelation of Hg^{2+} is due to its soft nature preferably binds with soft base $-NH_2$, –C=N, =N. The presence of interfering ions would affect the uptake of target ions but not appreciable, however the hydrogel can be used for the removal of Zn^{2+}, Cd^{2+} and Hg^{2+}. This hydrogel is better suited to Zn^{2+},

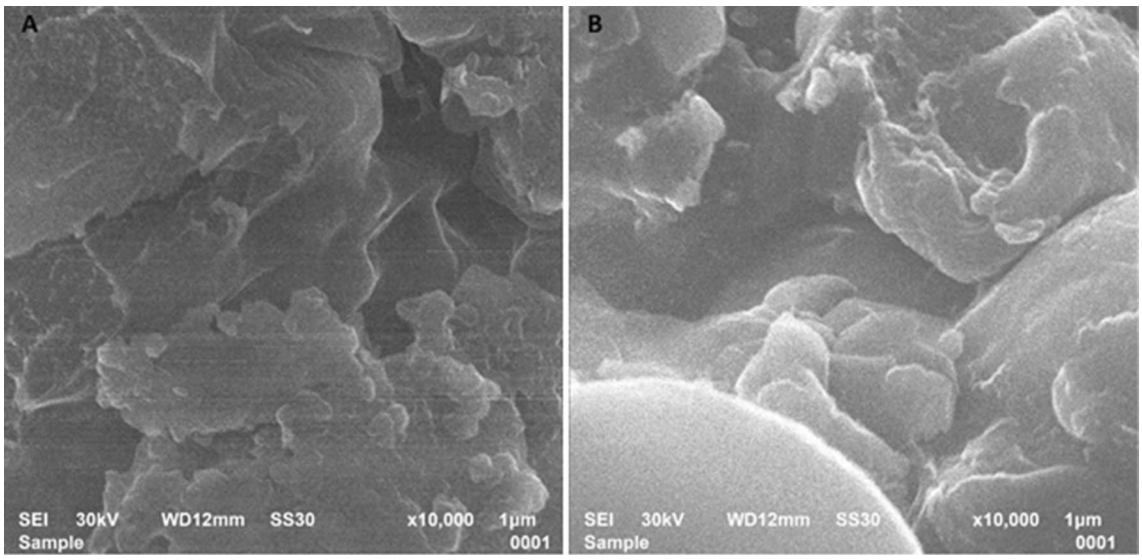


Fig. 4 Surface microphotographs of hydrogels

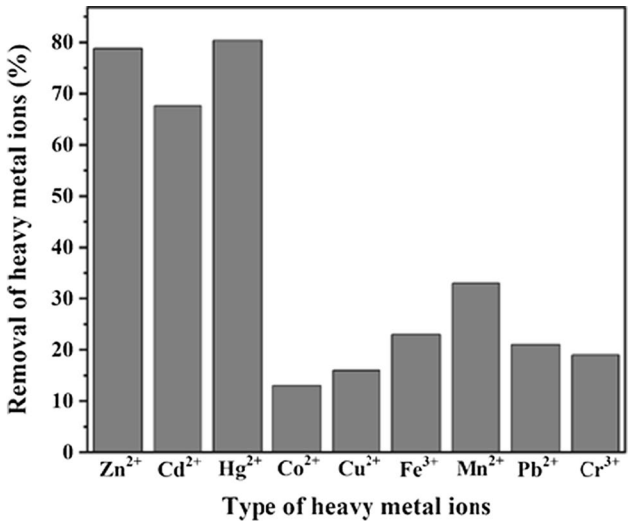


Fig. 5 Percentage removal of heavy metal ions using *meta*-BPDM PAM/CMG hydrogel (volume for each metal ion solution 10 mL, weight of hydrogel 82 mg, initial metal ion concentration C_o, 500 ppm at RT pH 6.5 and stirred magnetically at 250 rpm for 24 h)

Cd^{2+} and Hg^{2+} due to higher binding constants of embedded *meta*-BPDM ligands with these metal ions (Chauhan et al. 2020, 2021; Lone et al. 2019).

Individual heavy metal ion removal from aqueous solution (Non-Competitive Environment)

Non-competitive adsorption was studied using aqueous solution containing only one type of metal ion. The amount of Zn^{2+}, Cd^{2+} and Hg^{2+} removed by non-competitive adsorption is more than that can be achieved by competitive adsorption. The selective interaction between hydrogel and Zn^{2+}, Cd^{2+} and Hg^{2+} in water is more dominated in absence of other interfering metal ions. The aqueous solutions of Zn^{2+}, Cd^{2+} and Hg^{2+} of required concentrations were prepared using their chloride salts in deionized water. All experiments were carried out at pH 6.5 initially adjusted using HCL and NaOH, where the maximum adsorption was obtained. The weight of dry hydrogel pellets was 0.082 g uniformly in all batch experiments. The hydrogel was immersed in V 10 mL of aqueous solution with initial metal ion concentration C_o, 500 ppm at RT and stirred magnetically at 250 rpm for 24 h.

Metal absorbed or equilibrium adsorption capacity (mg/g),

$$q_e = \frac{V\left(C_o - C_e\right)}{W}$$

$$\text{Removal efficiency or Sorption } \% = \frac{\left(C_o - C_e\right)}{C_o} \times 100$$

$W = 82.0$ mg dosage of dried hydrogel, q_e (mg/g) and C_e (ppm) is the amount adsorbed and concentration left at equilibrium, respectively.

The absorption capacity and removal percentage of Zn^{2+}, Cd^{2+} and Hg^{2+} using *meta*-BPDM embedded PAM/CMG hydrogel (Table 2) was found to be higher than PAM/CMG hydrogel (blank) (Table 3).

Effect of initial metal ion concentration and contact time

The adsorption behaviour of hydrogel was studied in the batch experiment. Removal efficiency is significantly affected by the initial metal ion concentration, with increase in initial concentration of Zn^{2+}, Cd^{2+} and Hg^{2+} q_e also increases due to higher chances of collision between metal ions. Which results in utilization of all active adsorption sites on adsorbent. At high C_o the number of metal ions are higher than the available vacant sites in the hydrogel. Therefore, removal efficiency depends on the initial metal ion concentration. The maximum adsorption (q_e) at equilibrium was 47.67, 40.89 and 48.64 mg/g for Zn^{2+}, Cd^{2+} and Hg^{2+}, respectively, at RT, pH 6.5, dry hydrogel W 82 mg, volume of aqueous solution V 10 mL, Time 12 h, rpm 250, C_o (50, 150, 250, 300, 400, 500, 600,700 ppm), C_e = concentration at equilibrium and shown in Fig. 6a

$$q_e = \frac{V\left(C_o - C_e\right)}{W}$$

Contact time is a very important factor for the determination of equilibrium point, which can be used to study the kinetics of the adsorption process. The supernatants obtained after the hydrogel treatment were acidified and analysed by

Table 2 Uptake of zinc, cadmium and mercury by *meta*-BPDM PAM/CMG hydrogel

Metal ion	C_o (ppm)	C_e (ppm)	q_e (mg/g)	% Removal
Zn^{2+}	500	106	48.0	78.8
Cd^{2+}	500	162	41.21	67.6
Hg^{2+}	500	98	49.02	80.4

Table 3 Uptake of zinc, cadmium, and mercury by PAM/CMG (blank) hydrogel

Metal ion	C_o (ppm)	C_e (ppm)	q_e (mg/g)	% Removal
Zn^{2+}	500	265	28.43	47
Cd^{2+}	500	235	32.06	53
Hg^{2+}	500	280	26.62	26.62

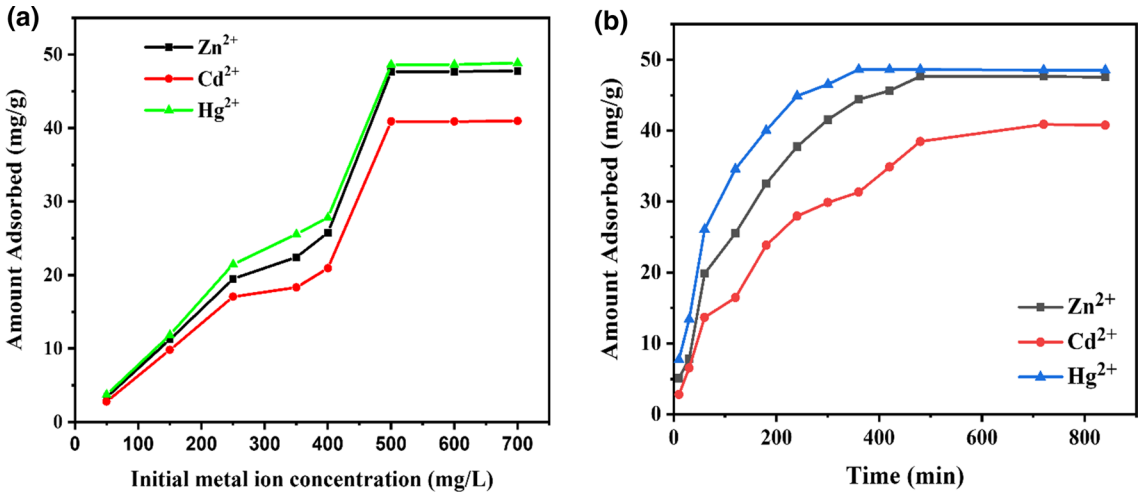


Fig. 6 Effect of initial metal ion concentration (**a**) and contact time (**b**) on adsorption of Zn^{2+}, Cd^{2+} and Hg^{2+} ions

ICP-MS as an independent verification at each time interval (10, 20, 30, 60, 120, 180, 240, 300, 360, 420, 480, 720 840 min). The decrease in concentration after each interval can be explained by the ability of *meta*-BPDM PAM/CMG hydrogel to bind with metal ions. These plots (Fig. 6a, b) with time indicate that the residual concentration of metal ions becomes asymptotic showing no appreciable change in the residual concentration of metal ions after 360 min in the case of mercury, 480 min in the case of zinc and 720 min in the case of cadmium. These represent the time at which equilibrium has been attained. q_t and C_t are the amount adsorbed (mg/g), residual concentration at time interval t, respectively (Fig. 6b).

$$q_t = \frac{V\left(C_o - C_t\right)}{W}$$

Effect of pH

The effect of pH plays the most important role in determining the parameters controlling removal of heavy metals from aqueous solutions. It affects the ionization or dissociation and precipitation of adsorbate and surface properties of the adsorbent. The effect pH is more prominent when both adsorbate and adsorbent are charged or polar species. The pH of the solution affects the adsorption efficiency by protonation of the polar functional group in the sorbent and precipitating metal ions as hydroxide or oxides. Hydrogen's ions may compete with metal ions for the active site on adsorbent at lower pH as well as adsorption sites that get protonated (Wang et al. 2016). The pH effect on the absorption of Zn^{2+}, Cd^{2+} and Hg^{2+} was studied in a pH range from 2 to 9. At pH > 6 the metal hydroxides would precipitate in aqueous solution that may affect the accuracy of results. The optimum pH was found to be 6.5 corresponds to maximum removal and above this metal hydroxide precipitation occurs. The observed results are displayed in Fig. 7.

Adsorption Isotherm studies

Adsorption isotherms represent how adsorbate and adsorbent molecules interact with each other, so are important in optimizing the utility of adsorbents. The obtained results were analysed by the Langmuir and Freundlich models. The Langmuir model assumes the adsorbent is having structurally homogeneous identical sorption sites with similar energies. The linear form of Langmuir equation is given by

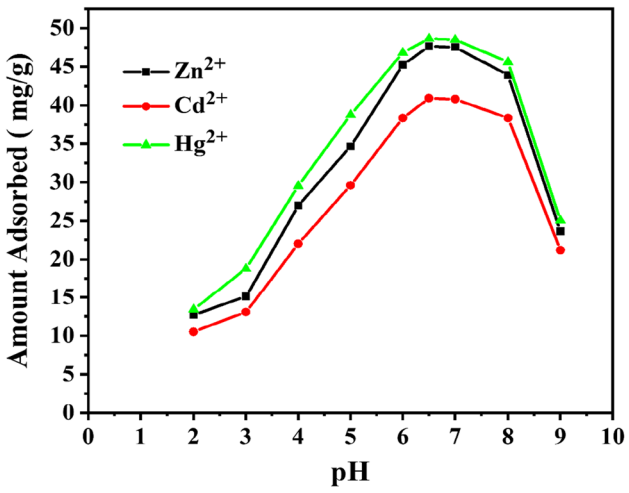


Fig. 7 Effect of pH on adsorption of zinc, cadmium and mercury. C_o = 500 ppm, W = 82.0 mg, V = 10 mL, RT, 250 rpm

$$\frac{1}{q_e} = \frac{1}{Q_o} + \frac{1}{bQ_oC_e}$$

Q_o and b are the Langmuir constants (Zhao et al. 2019; Ozay et al. 2009). The linearized form of the isotherm and high value of correlation coefficient (R^2) suggest that the adsorption follows the Langmuir model which indicates chemisorption (Fig. 8a).

The logarithmic equation of Freundlich model is given by

$$\log q_e = \log K_F + \left(\frac{1}{n}\right) \log C_e$$

where K_F and n are the adsorption capacity and adsorption intensity related Freundlich constants, respectively (Zhao et al. 2019; Meena et al. 2005; Gupta et al. 2003). Figure showing plots of $\log q_e$ against $\log C_e$ for the adsorption of zinc, cadmium and mercury. It is clearly visible that the data is not fitting very well in the Freundlich isotherm which indicates physisorption (Fig. 8b).

Adsorption kinetics

The pseudo-first and second-order model is the most common for solid liquid interaction and was applied to describe the sorption process. t is contact time, q_t and q_e are absorption capacity at time t and at equilibrium, k_1 and k_2 are rate constants for pseudo-first and second-order reactions. The Pseudo-first-order model considers rate of adsorption as controlled by mass transfer and diffusion whereas in case of the pseudo-second order it considers chemisorption is the rate controlling step (Gupta et al. 2003; Zhao et al. 2019).

$$\text{Absorption capacity, } q_t = \frac{V(C_o - C_t)}{W}$$

$$\log\left(q_e - q_t\right) = \log q_e - \frac{k_1}{2.303}t \quad \text{Pseudo-first order}$$

$$\frac{t}{q_t} = \frac{1}{k_2 q_e^2} + \frac{t}{q_e} \quad \text{Pseudo-second order}$$

The calculated q_e values of the pseudo-second-order match closely more with experimental as well as correlation coefficient (R^2) was higher (Table 4). Therefore, results obtained from a pseudo-second-order kinetic model better described the adsorption by hydrogel indicating chemisorption adsorption process Fig. 9a, b.

Removal of Zn^{2+} Cd^{2+} and Hg^{2+} from real wastewater

In order to assess the feasibility of the *meta*-BPDM PAM/CMG hydrogel for the removal of Zn^{2+}, Cd^{2+} and Hg^{2+} water. A real wastewater collection is collected from the vicinity of some industries nearby DTU, Delhi and other

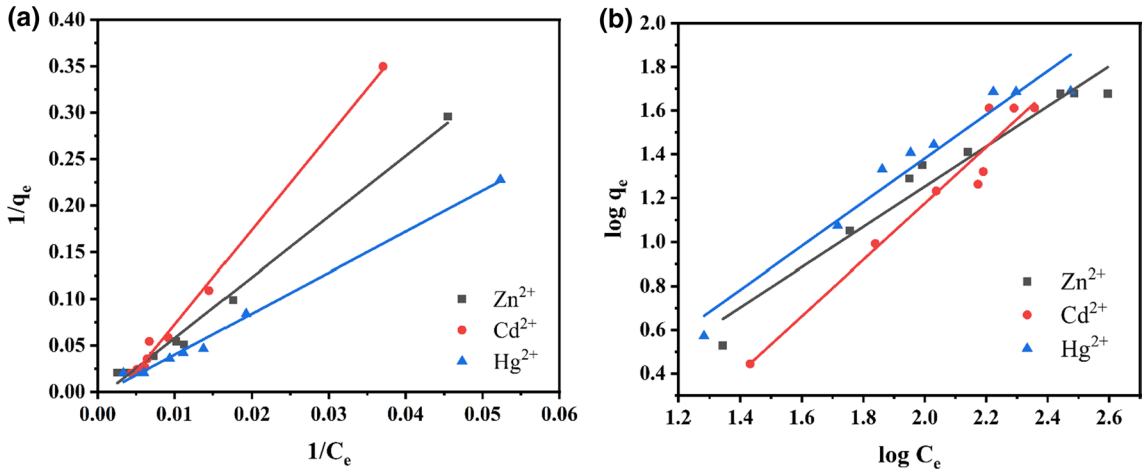


Fig. 8 Langmuir (**a**), Freundlich (**b**) isotherm model for Zn^{2+}, Cd^{2+} and Hg^{2+} ions adsorption

Table 4 Kinetic parameters for the adsorption of Zn^{2+}, Cd^{2+} and Hg^{2+} ions

Metal ion	Pseudo-first order			Pseudo-second order		
	k_1	q_{e1}	R^2	k_2	q_{e2}	R^2
Zn^{2+}	0.00663	43.74	0.9539	0.00011	51.42	0.9929
Cd^{2+}	0.00619	52.54	0.9432	0.00009	42.52	0.9950
Hg^{2+}	0.00838	25.02	0.7451	0.00028	53.09	0.9975

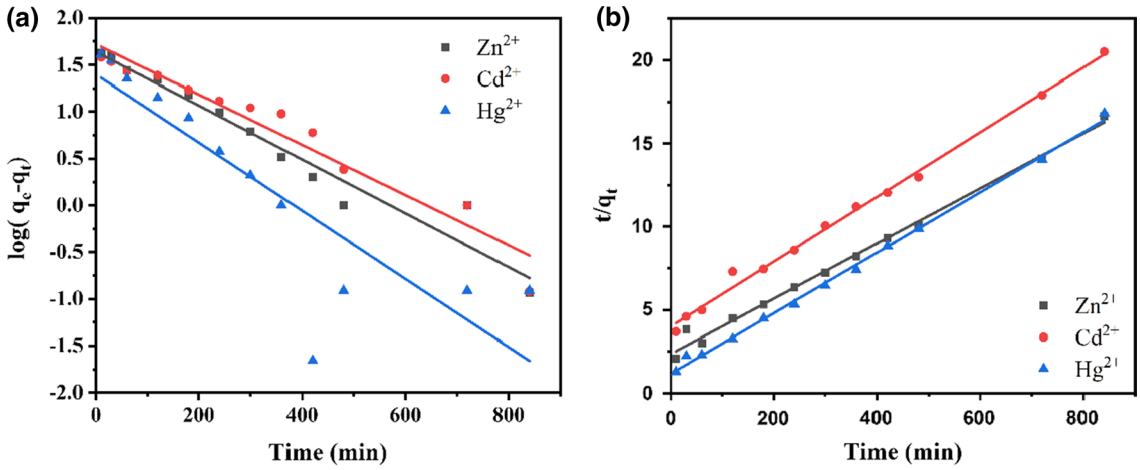


Fig. 9 Pseudo-first order (**a**) and pseudo-second order (**b**) kinetic model on Zn^{2+}, Cd^{2+} and Hg^{2+} ions adsorption

industrial areas of Delhi, India. ICP-MS showed very less amount of detectable Zn^{2+}, Cd^{2+} and Hg^{2+} (less than 3 ppm) in the collected waste water so chlorides salts of metal ions were added to simulate contaminated water (500 ppm). The water sample was transferred into the conical tubes at pH 6.5, RT, C_o 500 ppm, and a meta-BPDM PAM/CMG hydrogel-based sensor was added into the tubes, allowed to soak for 1 day. After soaking the supernatant solutions were collected and acidified using 1% HNO_3 for ICP-MS analysis. The concentrations in the supernatant were found to be decreased from 500 ppm to 113, 170, 102 ppm for Zn^{2+}, Cd^{2+} and Hg^{2+}, respectively, after the gel treatment. The results suggest that the hydrogel was capable of removing these metal ions from natural water sources (Najafi et al. 2011). Figure 10 depicts the hydrogel removal efficiency with respect to Zn^{2+}, Cd^{2+} and Hg^{2+} in water. Removal $\% = \frac{(C_o - C_e)}{C_o} \times 100$.

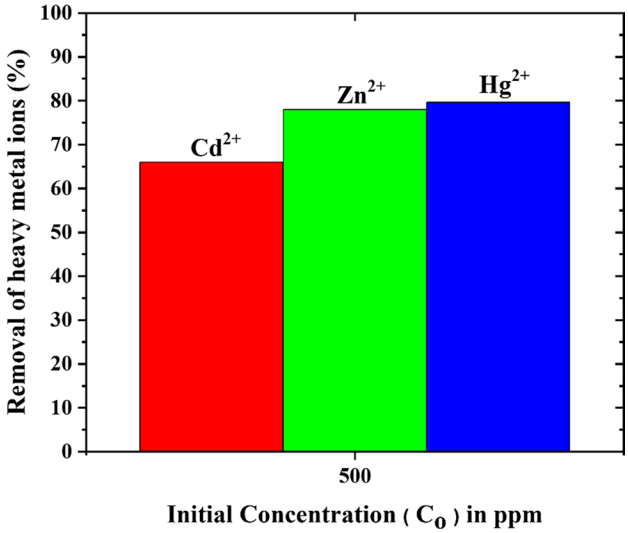


Fig. 10 Removal (%) of Zn^{2+}, Cd^{2+} and Hg^{2+} by hydrogel from industrial effluents at RT

Conclusion

The porous structure of modified copolymer-based hydrogel was beneficial for permeation of metal ions into it. The optimum pH was found to be 6.5 and Zn^{2+}, Cd^{2+} and Hg^{2+} achieved adsorption equilibrium within 12 h. The order of adsorption in the multicomponent system was found to be $Hg^{2+} > Zn^{2+} > Cd^{2+}$. The adsorption data fit very well to the Langmuir adsorption isotherm model and pseudo-second-order equation. It is found that the adsorption of Hg^{2+} is better than Zn^{2+} and Cd^{2+} under similar conditions due to difference in affinity or binding constants sequence $Hg^{2+} > Zn^{2+} > Cd^{2+}$ with the incorporated *meta*-BPDM in the PAM/CMG hydrogel. The obtained result establishes a suitable method with promising applications for efficient removal of Zn^{2+}, Cd^{2+} and Hg^{2+} from their aqueous solutions. Removal of heavy metal ions from aqueous medium using different formulations of hydrogel based on biopolymers such as raw wool, carboxymethyl tamarind etc. and *meta*-benziporphodimethenes is the current focus of our lab.

Acknowledgements Anil Kumar and Sudhir G. Warkar thanks to Delhi Technological University, Delhi, India for providing facilities for the study of the work and Deepti Chauhan thanks to The Council of Scientific & Industrial Research (CSIR), Delhi, India for Senior research fellowship and The Indian Agricultural Research Institute (The Pusa Institute) IARI, New Delhi, India for providing facilities of ICP-MS and AAS.

Author's contribution Deepti Chauhan: Writing—Original draft preparation, execution of the work, Methodology, Anil Kumar: Conceptualization, Methodology, Software, Sudhir G. Warkar: Reviewing, editing and conceptualization.

Funding No funding was received for conducting this study.

Availability of data and material Required data have been discussed in this manuscript itself.

Code availability Not applicable.

Declarations

Ethical approval This article does not contain any studies with animals performed by any of the authors.

Consent to participate All authors have verified and agreed for the publication of this manuscript.

Consent for publication All authors gave consent to submit this original research paper and obtained consent from the responsible institution before the work is submitted for the publication.

Conflict of interest The authors declare that they have no conflict of interest.

References

Abbas K, Znad H, Awual MR (2018) A ligand anchored conjugate adsorbent for effective mercury (II) detection and removal from aqueous media. Chem Eng J 334:432–443. https://doi.org/10.1016/j.cej.2017.10.054

Ahmad JU, Goni MA (2010) Heavy metal contamination in water, soil, and vegetables of the industrial areas in Dhaka, Bangladesh. Environ Monit Assess 166:347–357. https://doi.org/10.1007/s10661-009-1006-6

Atta A, Ismail HS, Mohamed HM, Mohamed ZM (2011) Acrylonitrile/acrylamidoxime/2-acrylamido-2-methylpropane sulfonic acid-based hydrogels: synthesis, characterization and their application in the removal of heavy metals. J Appl Polym Sci 122:999–1011. https://doi.org/10.1002/app.34245

Awual MR (2015) A novel facial composite adsorbent for enhanced copper (II) detection and removal from wastewater. Chem Eng J 266:368–375. https://doi.org/10.1016/j.cej.2014.12.094

Awual MR (2017) Novel nanocomposite materials for efficient and selective mercury ions capturing from wastewater. Chem Eng J 307:456–465. https://doi.org/10.1016/j.cej.2016.08.108

Bai C, Wang L, Zhu Z (2020) Adsorption of Cr (III) and Pb (II) by graphene oxide/alginate hydrogel membrane: characterization, adsorption kinetics, isotherm and thermodynamics studies. Int J Biol Macromol 147:898–910. https://doi.org/10.1016/j.ijbiomac.2019.09.249

Chauhan D, Kumar A, Warkar SG (2020) Synthesis, characterization and metal ions sensing applications of meta-benziporphodimethene-embedded polyacrylamide/carboxymethyl guargum polymeric hydrogels in water. Environ Technol (united Kingdom) 0:1–12. https://doi.org/10.1080/09593330.2020.1812730

Chauhan D, Kumar A, Warkar SG (2021) Modified polymeric hydrogels for the detection of Zn^{2+} in *E. coli* bacterial cells and Zn^{2+}, Cd^{2+} and Hg^{2+} in industrial effluents. Environ Technol (united Kingdom). https://doi.org/10.1080/09593330.2021.1928294

Dave N, Chan MY, Huang PJJ, Smith BD, Liu J (2010) Regenerable DNA-functionalized hydrogels for ultrasensitive, instrument-free mercury (II) detection and removal in water. J Am Chem Soc 132:12668–12673. https://doi.org/10.1021/ja106098j

Ding Y, Zhu W, Xu Y, Qian X (2015) A small molecular fluorescent sensor functionalized silica microsphere for detection and removal of mercury, cadmium, and lead ions in aqueous solutions. Sens Actuators B Chem 220:762–771. https://doi.org/10.1016/j.snb.2015.05.113

Gupta VK, Sharma S (2002) Removal of cadmium and zinc from aqueous solutions using red mud. Environ Sci Technol 36:3612–3617. https://doi.org/10.1021/es020010v

Gupta VK, Jain CK, Ali I, Sharma M, Saini VK (2003) Removal of cadmium and nickel from wastewater using bagasse fly ash—a sugar industry waste. Water Res 37:4038–4044. https://doi.org/10.1016/S0043-1354(03)00292-6

Gupta A, Rai DK, Pandey RS, Sharma B (2009) Analysis of some heavy metals in the riverine water, sediments and fish from river Ganges at Allahabad. Environ Monit Assess 157:449–458. https://doi.org/10.1007/s10661-008-0547-4

Jin L, Bai R (2002) Mechanisms of lead adsorption on chitosan/PVA hydrogel beads. Langmuir 18:9765–9770. https://doi.org/10.1021/la0259171

Kadirvelu K, Goel J, Rajagopal C (2008) Sorption of lead, mercury and cadmium ions in multi-component system using carbon aerogel as adsorbent. J Hazard Mater 153:502–507. https://doi.org/10.1016/j.jhazmat.2007.08.082

Karbarz M, Khalil AM, Wolowicz K, Kaniewska K, Romanski J, Stojek Z (2018) Efficient removal of cadmium and lead ions from water by hydrogels modified with cystine. J Environ Chem Eng 6:3962–3970. https://doi.org/10.1016/j.jece.2018.05.054

Kong C, Zhao X, Li Y, Yang S, Chen YM, Yang Z (2020) Ion-induced synthesis of alginate fibroid hydrogel for heavy metal ions removal. Front Chem 7:905. https://doi.org/10.3389/fchem.2019.00905

Liu H, Li S, Feng L, Hua Y, Cai Y, Yin M, Wan Y, Wang H (2020) A selective colorimetric and efficient removal strategy for mercury (II) using mesoporous silver-melamine nanocomposites synthesized by controlled supramolecular self-assembly. J Hazard Mater 388:121798. https://doi.org/10.1016/j.jhazmat.2019.121798

Lone S, Yoon DH, Lee H, Cheong IW (2019) Gelatin-chitosan hydrogel particles for efficient removal of Hg(ii) from wastewater. Environ Sci Water Res Technol 5:83–90. https://doi.org/10.1039/c8ew00678d

Lv Q, Hu X, Zhang X, Huang L, Liu Z, Sun G (2019) Highly efficient removal of trace metal ions by using poly (acrylic acid) hydrogel adsorbent. Mater Des 181:107934. https://doi.org/10.1016/j.matdes.2019.107934

Meena AK, Mishra GK, Rai PK, Rajagopal C, Nagar PN (2005) Removal of heavy metal ions from aqueous solutions using carbon aerogel as an adsorbent. J Hazard Mater 122:161–170. https://doi.org/10.1016/j.jhazmat.2005.03.024

Milosavljevic N, Debeljkovic A, Krušic MK, Milašinovic N, Üzüm ÖB, Karadag E (2014) Application of poly(acrlymide-co-sodium methacrylate) hydrogels in copper and cadmium removal from aqueous solution. Environ Prog Sustain Energy 33:824–834. https://doi.org/10.1002/ep.11854

Najafi M, Rostamian R, Rafati AA (2011) Chemically modified silica gel with thiol group as an adsorbent for retention of some toxic soft metal ions from water and industrial effluent. Chem Eng J 168:426–432. https://doi.org/10.1016/j.cej.2010.12.064

Namasivayam C, Ranganathan K (1995) Removal of pb(ii), cd(ii), ni(ii) and mixture of metal ions by adsorption onto 'waste' fe(iii)/cr(iu) hydroxide and fixed bed studies. Environ Technol (united Kingdom) 16:851–860. https://doi.org/10.1080/09593330.1995.9618282

Ozay O, Ekici S, Baran Y, Aktas N, Sahiner N (2009) Removal of toxic metal ions with magnetic hydrogels. Water Res 43:4403–4411. https://doi.org/10.1016/j.watres.2009.06.058

Perumal S, Atchudan R, Yoon DH, Joo J, Cheong IW (2019) Spherical chitosan/gelatin hydrogel particles for removal of multiple heavy metal ions from wastewater. Ind Eng Chem Res 58:9900–9907. https://doi.org/10.1021/acs.iecr.9b01298

Perumal S, Atchudan R, Yoon DH, Joo J, Cheong IW (2020) Graphene oxide-embedded chitosan/gelatin hydrogel particles for the adsorptions of multiple heavy metal ions. J Mater Sci 55:9354–9363. https://doi.org/10.1007/s10853-020-04651-1

Rocha CG, Zaia DAM, da Alfaya RV, S, Alfaya AA da S, (2009) Use of rice straw as biosorbent for removal of Cu (II), Zn (II), Cd (II) and Hg (II) ions in industrial effluents. J Hazard Mater 166:383–388. https://doi.org/10.1016/j.jhazmat.2008.11.074

Sahu S, Sahu UK, Patel RK (2019) Modified thorium oxide polyaniline core-shell nanocomposite and its application for the efficient removal of Cr (VI). J Chem Eng Data 64(3):1294–1304. https://doi.org/10.1021/acs.jced.8b01225

Sahu S, Mallik L, Patel RK et al (2020) Facile synthesis of poly o-toluidine modified lanthanum phosphate nanocomposite as a superior adsorbent for selective fluoride removal: a mechanistic and kinetic study. Chemosphere 252:126551. https://doi.org/10.1016/j.chemosphere.2020.126551

Sahu S, Bishoyi N, Patel RK (2021a) Cerium phosphate polypyrrole flower like nanocomposite: a recyclable adsorbent for removal of Cr (VI) by adsorption combined with in-situ chemical reduction. J Ind Eng Chem 99:55–67. https://doi.org/10.1016/j.jiec.2021.03.041

Sahu S, Bishoyi N, Sahu MK, Patel RK (2021b) Investigating the selectivity and interference behavior for detoxification of Cr (VI) using lanthanum phosphate polyaniline nanocomposite via adsorption-reduction mechanism. Chemosphere 278:130507. https://doi.org/10.1016/j.chemosphere.2021.130507

Sharaf MA, Arida HA, Sayed SA, Younis AA, Farag AB (2009) Separation and preconcentration of some heavy-metal ions using new chelating polymeric hydrogels. J Appl Polym Sci 113:1335–1344. https://doi.org/10.1002/app.30137

Wang X, Deng W, Xie Y, Wang C (2013) Selective removal of mercury ions using a chitosan-poly (vinyl alcohol) hydrogel adsorbent with three-dimensional network structure. Chem Eng J 228:232–242. https://doi.org/10.1016/j.cej.2013.04.104

Wang F, Zheng Y, Zhu Y, Wang A (2016) Preparation of chitosan-g-poly (vinylimidazole-co-2-acrylamido-2-methyl propane sulfonic acid) granular hydrogel for selective adsorption of Hg^{2+}. Water Air Soil Pollut. https://doi.org/10.1007/s11270-016-2812-4

Zhao B, Jiang H, Lin Z, Xu S, Xie J, Zhang A (2019) Preparation of acrylamide/acrylic acid cellulose hydrogels for the adsorption of heavy metal ions. Carbohydr Polym 224:115022. https://doi.org/10.1016/j.carbpol.2019.115022

Proceedings of the 6th International Conference on Communication and Electronics Systems (ICCES-2021)
IEEE Xplore Part Number: CFP21AWO-ART; ISBN: 978-0-7381-1405-7

Analog Analysis of Novel Ferroelectric-Dual Material Oxide Stack-Double Gate FET

Isha[1], Rishu Chaujar[2]
(e-mail: [1]isha24dahiya@gmail.com; [2]chaujar.rishu@dtu.ac.in)
[1,2]Department of Applied Physics, Delhi Technological University, Main Bawana Road, Delhi, 110042, India

***Abstract*—In this work, the analog performance for Ferroelectric-Dual Material Oxide Stack-Double Gate FET (Fe-DMOS-DG FET) has been assessed and its effectiveness in reducing short channel effects (SCEs) has been analyzed. The obtained results have been compared with the conventional MOSFET having the same channel length. This research work utilizes gate metal comprising of two different materials to have better control over the channel. To reduce the injection of hot carriers, in one case a ferroelectric layer is used on a thin SiO_2 interface layer, and in another case, a ferroelectric layer is used on an oxide stack consisting of high k-dielectric and SiO_2 as gate dielectric material. Parameters like ON current, OFF current, Ion/Ioff current ratio, Id vs. Vg plot, Transconductance vs. Vg plot, threshold voltage, and sub-threshold slope are considered. It has been recorded that the new device has less subthreshold current, resulting in a lower subthreshold slope when compared to bulk MOSFET. Also, the electric field and carrier concentration along the channel are investigated, and it has been found that the novel device enhances carrier transport in the channel. The silvaco ATLAS device simulation software is used to do the numerical calculations in order to validate the results.**

Keywords*— *Ferroelectric-Dual Material-Double Gate MOSFET (Fe-DM-DG MOSFET), Ferroelectric-Dual Material Oxide Stack-Double Gate FET (Fe-DMOS-DG FET), high k-dielectric, short channel effects.

I. INTRODUCTION AND BACKGROUND

The continuous reduction of device dimensions has enabled considerable technical advancement, resulting in denser, faster, and more practical integrated circuits. Developing devices that can perform massive computations with a limited footprint (in terms of size) has been a significant motivator. Most of this has been made possible by the fact that the size of ordinary MOSFET has diminished by a magnitude of around 20 over the last two decades, and the MOS transistors that are being fabricated these days have a channel of length less than 100 nm. [1]. Due to the reduction of the MOS transistor's size, there is an appreciable gate-source and gate-drain overlap causing edge effects leading to an increase in longitudinal field, which was negligible compared to the perpendicular field in long channel transistor [1]. Furthermore, short channel effects like injection of hot carriers into the gate oxide, velocity saturation of carriers, drain induced barrier lowering (DIBL), carrier multiplication due to impact ionization hence causing various leakage currents and surface scattering have increased due to downscaling of MOSFET. Therefore, the significant issue faced these days during transistor development is the reduction of short channel effects while improving the device's performance. For low-power and high-speed processes, the subthreshold swing and the DIBL should be as small as possible. However, at room temperature, the conventional MOSFET subthreshold swing has a theoretical limit of 60 mV/dec [11].

Various alternatives, including the multiple-gate structure and advanced technologies such as gate oxide engineering, gate work function engineering, and so on, are being used these days to suppress short channel effects. [2]-[3]. Out of the various gate geometries, double gate geometry is the most promising. Two gates offer better controllability from both sides of the channel, subsequently resulting in stronger switching signals [4]. The short channel effect suppression capability of such a Dual Material Gate (DMG) design has been discovered to be very effective [5]. In gate workfunction engineering, the gate metal of appropriate work function is used in such a way that charge carriers are accelerated rapidly in the channel [2]. The workfunction of the material at the source side is higher in a dual material gate configuration than the workfunction of the gate electrode near the source. Gate stack architecture is being incorporated to reduce high fringing fields and gate tunneling [9]. The most prominent motive in current literature and system design discussions is to improve switching speed and attain ultra-low power operation. The threshold voltage regulates the lower bound of the operating voltage, and in a traditional bulk MOSFET, lowering the threshold voltage increases the off-state current (Ioff), and hence the static power consumption [6]. To minimize power consumption, abrupt-turn-on devices are being extensively researched.

A. Metal Oxide Semiconductor Field Effect Transistors (MOSFETs)

2021 6th International Conference on Communication and Electronics Systems (ICCES) | 978-1-6654-3587-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICCES51350.2021.9489168

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:09:41 UTC from IEEE Xplore. Restrictions apply.

MOSFET consists of two heavily doped regions, source and drain, separated by a moderately substrate. The structure of an n-MOSFET is represented in figure 1. The voltage applied to

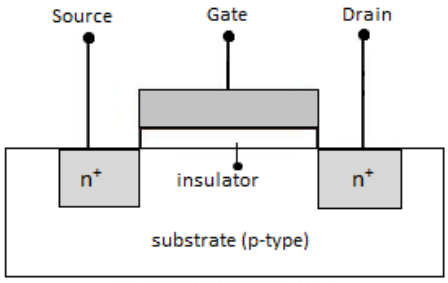


Figure 1. N-MOS structure

the gate terminal generates the channel by inversion of the semiconductor beneath the gate terminal. Inside the transistor, two electric fields govern its operation: the vertical electric field generated due to applied gate voltage and lateral electric field because of source-to-drain voltage. A dielectric material is needed to insulate the gate contact from the conducting channel; generally, SiO_2 is used. Since flow of current is limited to 10-100 Å at the semiconductor's surface, the bulk of the semiconductor, known as the substrate, is usually inactive [1].

The MOS transistor's operation is dependent on the MOS capacitor, which consists of gate metal contact, the substrate and the dielectric insulator that separates both of them. On applying voltage to the gate contact, charges are induced at the silicon-insulator interface in the semiconductor. For n-MOS transistor, the substrate is of p-type, therefore on applying Vg>0, holes start moving away from the Si-Insulator interface, forming a depletion layer that consists of immovable negative charge carriers. On further increasing the gate voltage, the depletion layer reaches the substrate further, and for a particular gate voltage, i.e. for the threshold voltage, electrons start injecting from the source to the drain, hence forming a channel.

B. Ferroelectric materials and FeFETs

Ferroelectric films are currently being studied in order to achieve ultra-low power operation in devices that incorporate them. All materials experience polarization in the presence of external electric field, which is proportional to the intensity of the applied field. However, ferroelectric dielectric crystals show a spontaneous electric polarization and have two stable polarized states, and an external electric field may be used to change the direction of polarization. These materials are ideal candidates for memory applications because of their switching efficiency. Digital bits, 0 or 1, are stored using the direction of spontaneous polarization, and the memory is non-volatile and electrically switchable. Typical hysteretic response of ferroelectric material is as shown figure 2. On the y-axis, the remnant polarization charges, denoted by PR+ and PR-, are

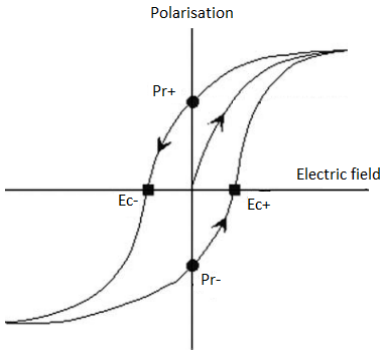


Figure 2. Hysteresis loop for ferroelectric material

shown, and on the x-axis, Ec+ and Ec- are marked, which represents the coercive fields.

The main difference between ferroelectric FETs and MOSFETs is found in the gate stack of the device. MOSFETs have an insulating dielectric layer between the gate and the channel, typically SiO_2, whereas FeFETs use ferroelectric materials. When doped with La, Sr, Si, Gd, Zr, Y, and Al, CMOS-compliant HfO_2-based films exhibit ferroelectric properties and are also lead-free [7]. However, there are specific challenges in interfacing Si and ferroelectrics, like intermixing of the two layers, and regardless of the application, all Fe FETs experience a substantial field drop across the low-k interface layer, which raises the power required for polarization reversal while decreasing lifetime [8]. Several buffer layer configurations and alternative gate stack layouts have been developed, like metal-ferroelectric-insulator-metal (MFIM), metal-ferroelectric-silicon (MFS), metal-ferroelectric-metal (MFM), and metal-ferroelectric-insulator-silicon (MFIS). The MFS structure is preferred to reduce the device's operating voltage; however, the immediate deposition of ferroelectric layer over silicon is complicated due to the chemical reaction [10]. Therefore, to shrink the chip's size for artificial intelligence operation, combinations of an oxide, a semiconductor, a thin ferroelectric material layer, and a fatigue-free electrode should be widely examined [10]. The approach to hysteresis-free FeFET architecture is aligning the positive capacitance of intrinsic MOSFET so as to stabilize the negative capacitance of ferroelectric; and experimentally ferroelectric-dielectric bilayer shows non-hysteretic electrical characteristics [12]. Also, ferroelectric based insulator design is found to improve the short channel immunity in conventional dual metal gate device [13] and use of high-k dielectric as interfacial layer improves endurance and retention in FeFET devices during operation at lower switching voltages [14]. Hence, it is anticipated that the use ferroelectric as insulator in dual gate structure along with an interface layer will improve device performance as compared to conventional MOSFET structure.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:09:41 UTC from IEEE Xplore. Restrictions apply.

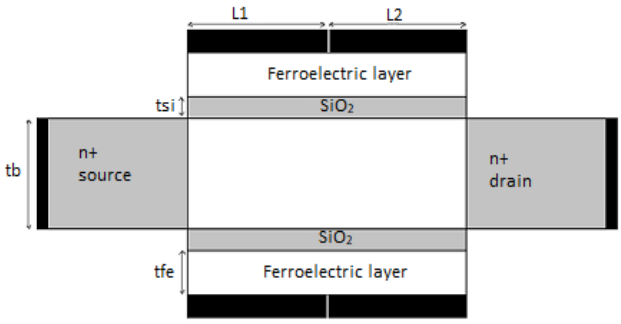


Figure 3. Schematic structure of Fe-DM-DG MOSFET

Device Description

Figure 3 depicts a cross-sectional illustration of a Ferroelectric-Dual Material-Double Gate MOSFET with ferroelectric material as a gate insulator and SiO_2 as an interface layer. The overall purpose of using an interface layer is to minimize inter-diffusion between the ferroelectric layer and the semiconductor. The thickness assumed for the ferroelectric and SiO_2 layer is 4nm (tfe) and 2nm (tsi), respectively. The simulations are performed for top and bottom gate of length (L1 + L2) 80nm and each gate are divided into two equal parts. Workfunction of gate regions near the source is taken to be 4.8eV and for the parts near the drain, it's 4.4eV. The thickness of the substrate (tb) used is 10nm with uniform doping of 10^{18} cm^{-3} and the source and drain regions have uniform doping of 10^{20} cm^{-3}. Figure 4 shows the cross-sectional structure of Ferroelectric-Dual Material Oxide Stack-Double Gate FET (Fe-DMOS-DG FET). Except for the gate oxide, which consists of a 1.5nm high-k material layer (tk) with dielectric constant 25, on a 1.5nm thick SiO_2 layer of dielectric constant 3.9, all other parameters are the same as above for Ferroelectric-Dual Material Oxide Stack-Double Gate FET. For the conventional MOSFET, work function of gate is taken to be equal to 4.6eV; substrate thickness is 30nm and the parameters like channel length and source/drain doping are the same as above.

To perform 2D numerical device simulation, mobility and recombination models included are standard concentration dependent mobility, Auger recombination model, Shockley-Read-Hall recombination with fixed carrier lifetimes, Bandgap Narrowing and parallel field mobility. For device simulation of FeFet devices, the Ferro model is incorporated to define the ferroelectric material [16].

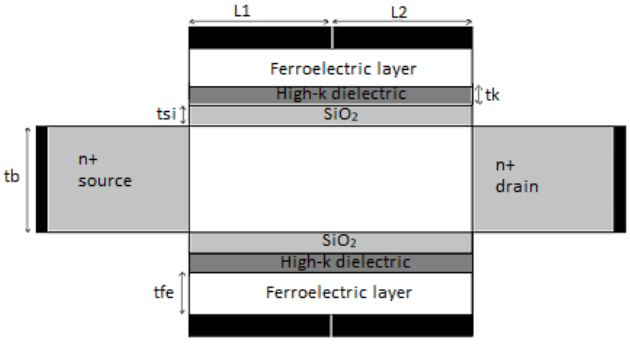


Figure 4. Schematic structure of Fe-DMOS-DG FET

II. Results and Discussion

This section presents the performance analysis of Ferroelectric-Dual Material-Double Gate MOSFET (Fe-DM-DG MOSFET) and Ferroelectric-Dual Material Oxide Stack-Double Gate FET (Fe-DMOS-DG FET) in comparison with conventional MOSFET (C-MOSFET). Figure 5 represents the transfer characteristics of the Fe-DM-DG MOSFET, Fe-DMOS-DG FET and C-MOSFET at Vds=0.5V. It can be seen that the threshold voltage for both the devices having ferroelectric material as gate insulator is higher as compared to the conventional MOSFET, which indicates the reduction of OFF state current.

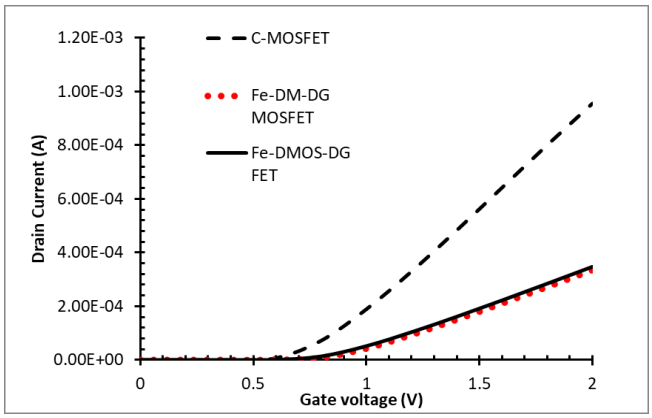


Figure 5. Drain Current vs. Gate Voltage (transfer characteristics) obtained for drain voltage=0.5V

Figure 6 shows transconductance vs. gate voltage plot for the three devices. Gate transconductance is expressed as the ratio of change in the output drain current to change in the applied gate voltage while maintaining the drain and substrate voltages constant. Transconductance is a crucial parameter since it evaluates device gain and controls device switching speed. Devices with high gain can operate at low gate voltages. The electric field in the channel increases as the device channel length decreases, resulting in velocity saturation of charge carriers, thereby limiting device current and transconductance. From the figure, transconductance increases at lower gate voltage; at higher gate voltage, it starts decreasing for C-MOSFET, however for the Fe-DM-DG MOSFET and Fe-DMOS-DG FET transconductance becomes independent of the gate voltage.

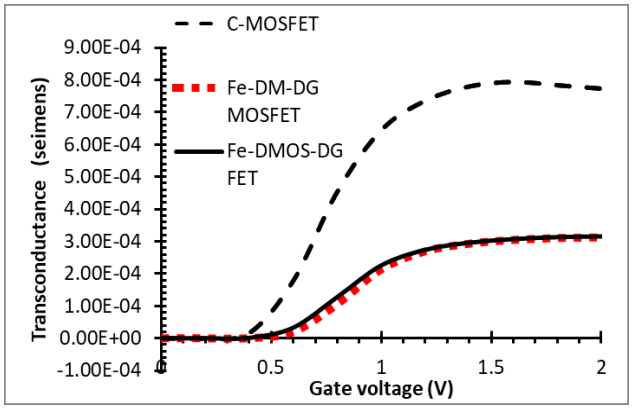


Figure 6. Transconductance vs. Gate voltage

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:09:41 UTC from IEEE Xplore. Restrictions apply.

Table 1. Electrical characteristics extracted at Vds=0.5V

Model	Sub-Threshold slope (V/dec)	Vth	Ion(Vg=1V)	Ioff(Vg=0V)	Ion/Ioff	DIBL
C-MOSFET	0.0929308	0.541862	0.000189	3.56E-10	530834	0.473247
Fe-DM-DG MOSFET	0.0785244	0.679169	4.24E-05	6.01E-15	7.06E+09	0.494781
Fe-DMOS-DG FET	0.0853542	0.648264	5.17E-05	1.08E-13	4.79E+08	0.504489

Table 1 compares electrical characteristics of C-MOSFET, Fe-DM-DG MOSFET and Fe-DMOS-DG FET for drain voltage (Vds) equal to 0.5V. The electrical features extracted are sub-threshold slope, threshold voltage (Vth), on and off current, the ratio of Ion and Ioff and DIBL. Fe-DM-DG MOSFET gives the lowest off-state current with highest ratio of Ion and Ioff. Both the ferroelectric devices have lower sub-threshold slope values than the conventional MOSFET; and threshold voltage increases implying better short channel immunity. However, a minimal increase in values of DIBL can be seen for both the new devices, which is defined as $|\Delta Vth|/|\Delta Vds|$. The values of Vth at Vds= 0.05V and 1.0V are taken to derive ΔVth in this study.

Figure 7 shows the variation of electric field in the channel for C-MOSFET, Fe-DM-DG MOSFET and Fe-DMOS-DG FET. The reduction of channel length leads to various leakage currents due to the high electric field in the channel. It can be seen that the electric field along the channel in C-MOSFET is higher as compared to the Fe-DM-DG MOSFET and Fe-DMOS-DG FET. The high value of field towards the drain terminal of the device, which leads to pronounced hot carrier effects in conventional MOSFET, has been decreased with the introduction of a ferroelectric layer as gate insulator. No significant impact of introducing an oxide stack in place of the SiO_2 layer in Fe-DM-DG MOSFET is observed on the electric field variation along the channel.

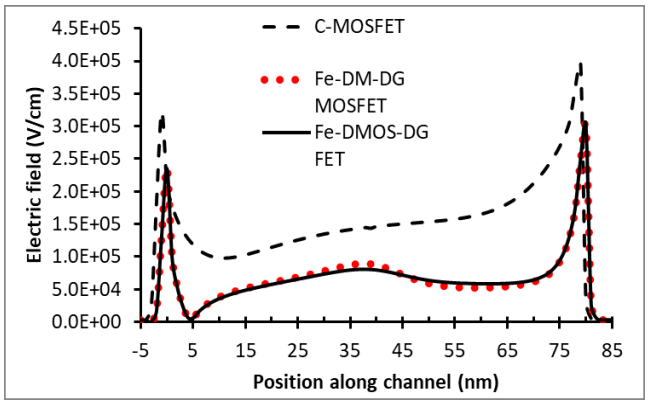


Figure 7. Variation of electric field along the channel for C-MOSFET, Fe-DM-DG MOSFET and Fe-DMOS-DG FET

Figure 8 compares the variation of electron concentration along the channel for all three devices. It is observed that the carriers are distributed non-uniformly in the C-MOSFET, which leads to mobility (μ) degradation, and the exact thickness of the conducting channel could be challenging to determine. Meanwhile, the gate controllability over the carriers in the channel would be reduced, and high carrier concentration towards the drain side could increase the hot carrier effects in the device. The introduction of ferroelectric layer has improved the electron density towards the source end, and the carrier concentration becomes more uniform along the channel compared to the C-MOSFET. However, variation of electron concentration is the same for Fe-DM-DG MOSFET and Fe-DMOS-DG FET.

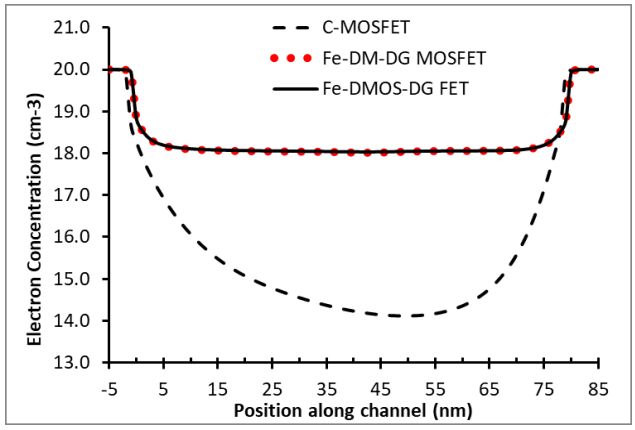


Figure 8. Variation of electron concentration along the channel for C-MOSFET, Fe-DM-DG MOSFET and Fe-DMOS-DG FET

Figure 9 shows potential distribution along the channel for three of the devices. In C-MOSFET, there is a formation of potential well which spans whole of the channel length, whereas in both of the ferroelectric devices, a potential barrier near the source is formed because of negative capacitance

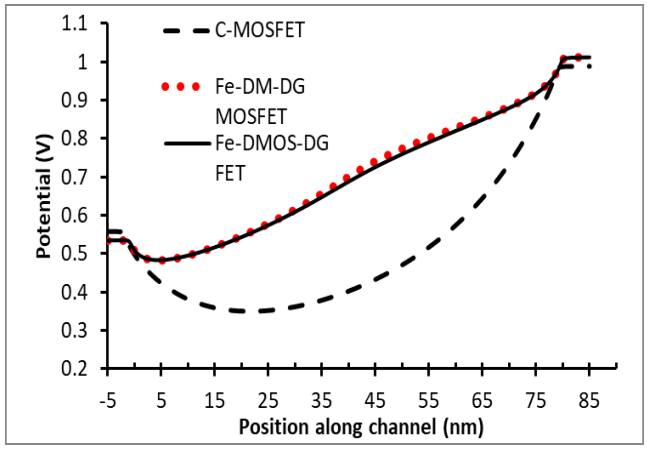


Figure 9. Variation of potential (V) along the channel

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:09:41 UTC from IEEE Xplore. Restrictions apply.

effect of ferroelectric layer. The higher potential towards source in Fe-DM-DG MOSFET and Fe-DMOS-DG FET causes injection of carrier from the source into channel at higher speed as compared to C-MOSFET. This results in improved short channel immunity.

III. CONCLUSION

A two-dimensional analytical model for Ferroelectric-Dual Material-Double Gate MOSFET (Fe-DM-DG MOSFET) and Ferroelectric-Dual Material Oxide Stack-Double Gate FET (Fe-DMOS-DG FET) has been developed to diminish short-channel effects (SCEs) and hot carrier effects (HCEs) in nano-scale bulk MOSFETs, as well as to improve device reliability. Fe-DM-DG MOSFET improves nearly all of the parameters such as off current, sub-threshold slope, variation of carrier concentration along channel when compared to conventional short channel MOSFET; and is found to be effective for reduction of HCEs. Further, the parameters like variation of electric field, potential and carrier concentration along the channel obtained for Fe-DM-DG MOSFET and Fe-DMOS-DG FET are nearly identical. Thus, it has been illustrated that incorporating a ferroelectric layer along with the SiO_2 layer as insulator for dual metal gate electrode configuration improves hot carrier reliability and short channel immunity, resulting in improved performance and better carrier transport efficiency when compared to conventional bulk MOSFET design.

Dual Gate FETs have assuring future in VLSI and IC design as they have notably small structures. Other advantages include less power consumption, high reliability, and mostly nano-scale size, making them very valuable for CMOS circuits. Incorporating ferroelectric material and Dual Gate structure in FETs can provide promising results in fast programming, high ON/OFF ratio, low power dissipation, and other applications, which makes them suitable to be used as an electronics synapse.

References

[1] Arora, Narain, ed. *Mosfet modeling for VLSI simulation: theory and practice*. World Scientific, 2007.

[2] Oh, Sang-Hyun, Don Monroe, and J. M. Hergenrother. "Analytic description of short-channel effects in fully-depleted double-gate and cylindrical, surrounding-gate MOSFETs." *IEEE electron device letters* 21, no. 9 (2000): 445-447.

[3] Long, Wei, and Ken K. Chin. "Dual material gate field effect transistor (DMGFET)." In *International Electron Devices Meeting. IEDM Technical Digest*, pp. 549-552. IEEE, 1997.

[4] Jung, Hak Kee, and Sima Dimitrijev. "Analysis of subthreshold carrier transport for ultimate DGMOSFET." *IEEE transactions on electron devices* 53, no. 4 (2006): 685-691.

[5] Reddy, G. Venkateshwar, and M. Jagadesh Kumar. "A new dual-material double-gate (DMDG) nanoscale SOI MOSFET-two-dimensional analytical modeling and simulation." *IEEE Transactions on Nanotechnology* 4, no. 2 (2005): 260-268.

[6] Cristoloveanu, Sorin, Jing Wan, and Alexander Zaslavsky. "A review of sharp-switching devices for ultra-low power applications." *IEEE Journal of the Electron Devices Society* 4, no. 5 (2016): 215-226.

[7] Upadhyay, Navnidhi K., Hao Jiang, Zhongrui Wang, Shiva Asapu, Qiangfei Xia, and J. Joshua Yang. "Emerging memory devices for neuromorphic computing." *Advanced Materials Technologies* 4, no. 4 (2019): 1800589.

[8] Winkler, Felix. "Through Silicon Via Field-Effect Transistor with Hafnia-based Ferroelectrics and the Doping of Silicon by Gallium Implantation Utilizing a Focused Ion Beam System." (2020).

[9] Gupta, Neha, and Rishu Chaujar. "Optimization of high-k and gate metal work function for improved analog and intermodulation performance of Gate Stack (GS)-GEWE-SiNW MOSFET." *Superlattices and Microstructures* 97 (2016): 630-641.

[10] Oh, S., H. Hwang, and I. K. Yoo. "Ferroelectric materials for neuromorphic computing." *APL Materials* 7, no. 9 (2019): 091109.

[11] Madan, Jaya, R. S. Gupta, and Rishu Chaujar. "Performance investigation of heterogeneous gate dielectric-gate metal engineered–gate all around-tunnel FET for RF applications." *Microsystem Technologies* 23, no. 9 (2017): 4081-4090.

[12] Li, Yang, Yong Lian, Kui Yao, and Ganesh S. Samudra. "Evaluation and optimization of short channel ferroelectric MOSFET for low power circuit application with BSIM4 and Landau theory." Solid-State Electronics 114 (2015): 17-22.

[13] Mehta, Hema, and Harsupreet Kaur. "Subthreshold analytical model for dual-material double gate ferroelectric field effect transistor (DMGFeFET)." Semiconductor Science and Technology 34, no. 6 (2019): 065008.

[14] Ali, T., P. Polakowski, S. Riedel, T. Büttner, T. Kämpfe, M. Rudolph, B. Pätzold et al. "High endurance ferroelectric hafnium oxide-based FeFET memory without retention penalty." *IEEE Transactions on Electron Devices* 65, no. 9 (2018): 3769-3774.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:09:41 UTC from IEEE Xplore. Restrictions apply.

2021 Fourth International Conference on Computational Intelligence and Communication Technologies (CCICT) | 978-1-6654-2392-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/CCICT53244.2021.00014

Analysis of Automated text generation using Deep learning

Dr. Manoj Kumar
Dept. of Computer Science and Engineering
Delhi Technological University
New Delhi, India
mkumarg@dce.ac.in

Abhishek Singh
Dept. of Computer Science and Engineering
Delhi Technological University
New Delhi, India
abhisheksanu26@gmail.com

Arnav Kumar
Dept. of Computer Science and Engineering
Delhi Technological University
New Delhi, India
arnavkumar1999@gmail.com

Ankit Kumar
Dept. of Computer Science and Engineering
Delhi Technological University
New Delhi, India
ankitkumar50417@gmail.com

***Abstract*—**

A chatbot is a computer program that can converse with humans using artificial intelligence in messaging platforms. The goal of the project is to use and optimize deep learning techniques for making an efficient chat bot. Current chatbots are developed using rule-based techniques, basic machine learning algorithms or retrieval-based techniques which don't generate good results.in this paper, we will be comparing performance of three chatbots build by using RNN, GRU and LSTM. The conversation chatbots are used by different companies, government organizations and more.

***Index Terms*— Natural Language Processing, Deep learning, Chatbots, Gated Recurrent Units, Long short-term memory.**

I. INTRODUCTION

Text generation is task of giving our sentences as input and getting a response from the machine/computer. It is done using deep learning algorithms which are subset of Neural networks and give the best result and accuracy. It is interesting problem in the field of Natural Language Processing. We only have a supervised learning approach for text generation and that is present in sequence to sequence models of the RNNs. This Project is intended to apply analysis different types of sequence to sequence models for generating text which should be meaningful to the user (not random text). We applied three methods i.e Simple Recurrent Neural Network, LSTM and GRU and compared them to find out the difference in their performance we have also looked other research papers for the results on basic RNN and LSTM. We used the Cornell Movies Dataset for this paper. We tried to figure out the using existing learning algorithm like GRU how can we improve it to get it on par with LSTM Ease of Use

II. HISTORY AND RELATED WORK

A. Early Approach

They are many open source bot API on the internet, providing an easy way to developers to create bots. Bots like Mitsuku were among the smartest bot created. But in the beginning it was like a pattern matching the developer used to write a pattern a template such that when matching ours it responds with one of the templates. Rule Based make it easy for everyone to create a bot but it is quite inconvenient hard to make bot that answers complex queries. There are many approaches in making a chat box some are AI approaches and other are Non-AI approaches in this paper we will be talking about the AI approaches i.e. Retrieval Based and Generative Models.in the title or heads unless they are unavoidable.

B. Intelligent Models

1) ***Retrieval Based*****:** Retrieval Based Chat bot has a collection of predefined responses and the bot is trained to give the best fitting response from the collection of responses. In this model all the responses are to manually entered by a human. It doesn't generate any new responses so we don't need to worry about grammar. It usually tends to make mistake as it has a limited set of responses due to which it tends to be rigid responses appears to non -human. It's basic job was to select the most optimal response from the pool of predefined responses.

2) ***Generative Model:*** This model doesn't uses any predefined responses instead it generate response itself word by word due to which they are prone to grammatical errors. This method is used for developing a smart more human like bot that are quite advanced in nature.Its responses are still unpredictable the algorithms are quite complex. we will be comparing three of the generative model algorithms further in the paper.

C. Related Work

Previously, Seq2Seq AI chat bots were developed using encoder-decoder architecture. The encoder-decoder was built using a modified Recurrent Neural Network know as Long Short Term Memory cell units. Seq2Seq model predicts utterances unit by unit not considering what could occur in future which also leads to many frequent responses like "I don't get what you are saying" which was also because of the dataset having such sentences to avoid these issues Deep Reinforcement learning is used which is beyond the scope of this paper. Adam optimizer were also used as computes individual adaptive learning rates for different parameters from estimates of first and second moments of the gradient. It was also selected as it is a combination of best properties of AdaGrad and RMSProp [8]. Bidirectional LSTM were also used before in encoder side and attention mechanism used in decoder side to improve model

978-1-6654-2392-2/21/$31.00 ©2021 IEEE
DOI 10.1109/CCICT53244.2021.00014

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:05:21 UTC from IEEE Xplore. Restrictions apply.

performance. The advantage that the Bidirectional LSTM is that it works like two independent Recurrent Neural Network put together enabling the network to have both forward backward information of the sequence at every step of the word generation [10] i.e. one transfer information from past to future the other transfer it from future to past as it runs backwards it preserve information from the future and using the two hidden states combined you are able in any point in time to preserve information from both past and future.

III. PRELIMINARIES

Here we will discuss some various deep learning techniques used for implementing chatbots.

A. Text Generation

Text generation comes under the branch of Natural Language processing where we use different deep learning algorithms and methods like sequence-to-sequence model with Recurrent neural networks, Gated recurrent units and long short-term memory respectively [1]. We can also different classification techniques to categorize words and generate sentences like Naive Bayes classifier.

B. Recurrent Neural Network

Special Type of Neural Network designed for natural language processing. Here we have our word embeddings or the bag of words model given as input to the recurrent neural network, then they'll produce an output based on the model (for our case it is text generation). Although the structure of this Neural Networks may seem complex it is not very good when we have different contexts. RNN normally will find it difficult to distinguish after the plural or singular words gets past as it doesn't have any memory to store the form of sentence.

C. Long Short Term Memory

Short-term, long-term memory network is a type of recurrent neural network (RNN). LSTM (Sepp Hochreiter et al [4]) are ideal for learning, processing and classifying sequential data. Its networks aim to overcome the problem of fading gradients by using gateways to selectively keep the information up to date and forget irrelevant details [3]. Reduced sensitivity to time intervals makes LSTM networks better for sequential data analysis than individual RNNs.

1) Architecture: The architecture of the LSTM (Sepp Hochreiter et al [4]) block is shown below. An LSTM unit typically has a storage cell, intake valve, exhaust valve, and bypass valve in addition to the hidden state of the RNN. The entry gate weights and offsets control how far the new value enters the cell. Likewise, the weights and offsets for the Oblivion gate and output control the extent to which the value remains in the cell and the extent to which the value in the cell is used to calculate the output activation of the LSTM-Block respectively [4].

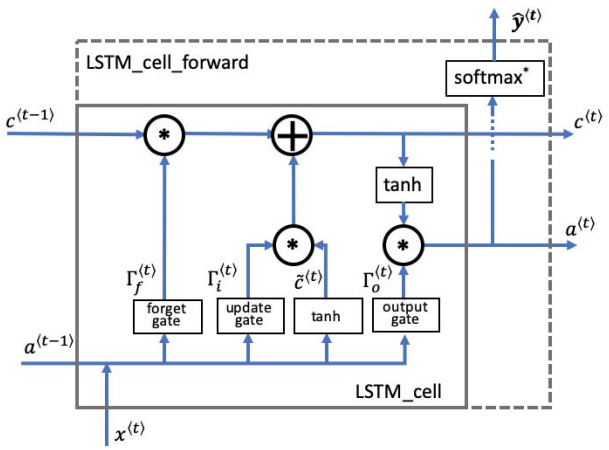


Fig. 1: LSTM unit

The Important part of LSTM is its cell state. The line passing through all the LSTM Units. Cell state acts as a memory unit to carry forward certain aspects of our sentence [3] as we know plural words cause issues for basic RNN, LSTM doesn't suffer from that.

$$i_t = \sigma(W_{ii}x_t + b_{ii} + W_{hi}h_{t-1} + b_{hi}) \quad (1)$$

$$f_t = \sigma(W_{if}x_t + b_{if} + W_{hf}h_{t-1} + b_{hf}) \quad (2)$$

$$g_t = tanh(W_{ig}x_t + b_{ig} + W_{hg}h_{t-1} + b_{hg}) \quad (3)$$

$$o_t = \sigma(W_{io}x_t + b_{io} + W_{ho}h_{t-1} + b_{ho}) \quad (4)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (5)$$

$$h_t = o_t \odot tanh(c_t) \quad (6)$$

where h_t is hidden state at time t, c_t is the cell state at time t , x_t is the input at time t, h_{t-1} is the hidden state of the layer at time t-1 or the initial hidden state at time o, and i_t, f_t , g_t , o_t are the input, forget, cell and output gates respectively [4] . σ is the sigmoid function, and $\odot$ is the Hadamard product. In the multi layer LSTM, the input $x_t^{(l)}$ of the lth layer (l>=2) is the hidden layer $h_t^{(l-1)}$ of the previous layer multiplied by dropout $\delta_t^{(l-1)}$ where each $\delta_t^{(l-1)}$ is a Bernoulli random variable which is 0 with probability dropout

D. Gated Recurrent Unit

The gated recurrent block (GRU) is a gating mechanism in recurrent neural networks (RNN), similar to the LSTM block, but without the outlet value. GRU (Jun young Chung et al [5]) tries to solve the problem of the vanishing gradient that can occur with standard recurrent neural networks. It can be thought of as a type of LSTM block as both are similar in design and in some cases give similar results.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:05:21 UTC from IEEE Xplore. Restrictions apply.

GRU can solve the disappearing gradient problem with the update and reset gate. The update gateway manages information flowing into memory, and the reset gateway contains information flowing out of memory [5]. GRU is a useful mechanism to solve the problem of vanishing gradients in recurrent neural networks. The vanishing gradient problem occurs in machine learning when the gradient becomes vanishingly small, which prevents the weight from changing its value.

1) ***Architecture:*** The architecture of the GRU block is shown below. An GRU unit typically has a forget, update gate and reset gate in addition to the hidden state of the RNN it lacks the output gates. The update gateway and the reset valve are two vectors that decide which information is transmitted to the output. We can train this to remember information from the past or to delete information unrelated to the forecast. [5]

2) ***Advantages:***

- GRU perform better than LSTM when working with small data sets.
- GRU is Computationally less expensive as compared to LSTM as we have fewer matrices to compute. it Works well with Smaller Datasets.

3) ***Disadvantages:***

- Performs worse in case of larger datasets as compared to LSTM.
- Lesser bleu score than LSTM.

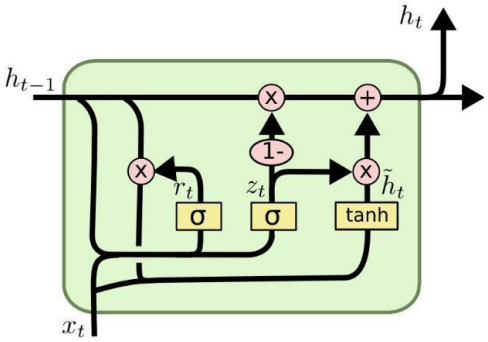


Fig. 2: GRU unit

4) ***Equations:***

$$z_t = \sigma_g(W_z x_t + U_z h_{t-1} + b_z) \quad (7)$$

$$r_t = \sigma_g(W_r x_t + U_r h_{t-1} + b_r) \quad (8)$$

$$\hat{h}_t = \phi_h(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h) \quad (9)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \hat{h}_t \quad (10)$$

E. Encoder Decoder Model with Attention layer

1) Encoder: It is a Stack of a recurrent units (of whichever model being used like LSTM or GRU) which takes element of input sentence or sequence and pass on its information to further layers Equation for hidden states vector [6], [7].

$$h_t = f(W^{hh} h_t - 1 + W^{hx} x_t) \quad (11)$$

This equation denotes final hidden layer vector produced by the encoder this helps in combining all the information from the input and then providing it to decoder for respective prediction or in this case generating text. This vector is directly given as first input layer of the decoder [6], [7]. hidden layer and memory value of each unit is given to the attention layer [2].

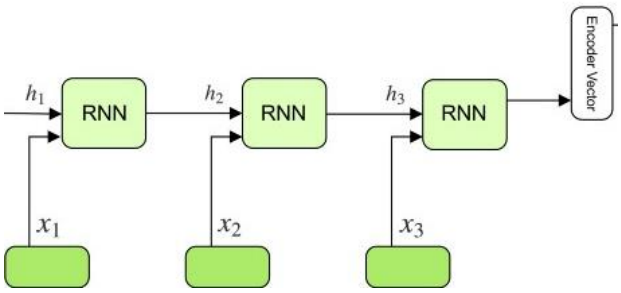


Fig. 3: Encoder in Seq2Seq model

2) Attention Layer: Computes 4 important elements which are score of each encoder state, attention weights using SoftMax function over the aforementioned scores, Context vector by multiplying hidden layer vector and attention weights and concatenate them with previous time step and send this to the decoder along with last hidden layer vector [2]. We have 2 types of attention models which are as follows:

1. **Global Attention layer:** Here all encoder states are considered for obtaining the context vector

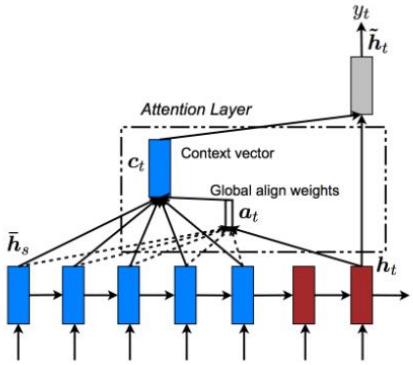


Fig. 4: Global attention layer - l – at each time step t, the model infers a variable-length alignment weight vector at based on the current target state h_t and all source states h_s. A global context vector c_t is then computed as the weighted average, according to at, over all the source state (Minh-Thang Luong et al [2]).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:05:21 UTC from IEEE Xplore. Restrictions apply.

2. Local Attention layer: Global attention has a few drawbacks like it attends all words on the encoder side for all target words which could potentially be computationally expensive. Here local attention comes into action where we have a small subset of source positions per target word [2]. Equations for attention weights and context vector are referred from Minh-Thang Luong et al [2].

3) Decoder: Decoder is used to generate text in case of Seq2Seq model where it takes the context vector from the attention layer and last hidden vector from the encoder and gives output accordingly.

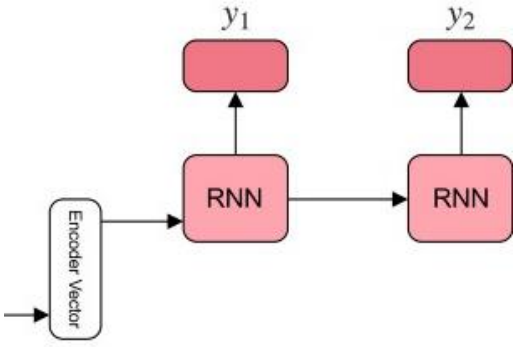


Fig. 5: Decoder in Seq2Seq model

IV. METHODOLOGY

Here we will judge performance of aforementioned algorithms and we will use our model in the mix as well.

Algorithm: Deep Neural Network, Recurrent Neural Network.

Main Technique: We have used the Seq2Seq model with attention layer [2] with GRU for this paper We will be comparing our outputs with paper which has implemented same model on same dataset using LSTM.

frameworks used: We have used various deep learning frameworks like PyTorch for respective implementation.

Optimization technique: For this paper we have used Adam optimization technique to get the best results which we can.

V. CONFIGURATION

We have used 3 different algorithms 2 of which are implemented by use other is from the base paper [9].

TABLE I: Configuration Table

Parameters	*LSTM [9]*	*RNN*	*GRU*
batch size	32	128	128
embedding size	1024	1024	1024
RNN Size	1024	1000	1250
learningrate	0.001	0.0001	0.00001
epochs	50	32	46
minlearningrate	0.0001	0.00001	0.000001
learning rate decay	0.9	0.9	0.9

VI. TRAINING MODEL

We trained the model for 2 configurations as we discussed above using Adam Optimization which can vary the learning
rate using learning rate decay technique [8]. we had around 82000 sentences in Cornell movies dataset after tokenization. We used Bidirectional GRU in encoder side and decoder is made with help of Luong attention layer [2].

VII. HARDWARE SPECIFICATION

Training on google colab for GRU and RNN Basic Seq2seq and my local machine configuration is stated below

TABLE II: Specification

Hardware	Specification
Processor	Core-i5 6400 @3.3 GHz
RAM	DDR4 16GB @ 2400MHz
GPU	Radeon RX 5700 XT 8GB

VIII. COMPARISON

We Chose a LSTM based research paper [9] for base of this project in order to compare performance of this model with LSTM based implementation to see our model is any good or not. As for Other GRU implementations we saw different models which were basic implementations so it would not be a fare comparison for our work. All in all, we compared this model to the basic GRU implementations and by observing the results we have seen certainly there is an improvement however we were unable to include the comparisons because of limited resources. We have mentioned all the parameters we tweaked in order to obtain our results in the configuration section of this paper.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:05:21 UTC from IEEE Xplore. Restrictions apply.

IX. RESULT

Here a test is given from all 3 models after seeing training time and responses GRU is the best option for this dataset to create chatbot but if the dataset is more complex then LSTM is the best choice but work is still being done on GRU so it has scope to improve even further.

Human :hey
Bot :hi
Human :what doing
Bot :sleep
Human :when will you be okay
Bot :just few days
Human :where are you from
Bot :southern California
Human: you want to come with me?
Bot: how is that for you?
Human :how much do you like me
Bot :you are not cold mostly
Human :who am i
Bot :doctor Livingston
Human :what's your hobby
Bot :eat
Human :wow you are best
Bot :I have done
Human :really well
Bot :goodbye

(a) LSTM [9] conversation

```
> hey
Bot: there
> what ypu doing
Bot: there
> when will you be okay
Bot: i m sorry .
> where are you from
Bot: i m sorry .
> you want to come with me
Bot: i m not .
> how much do you like me
Bot: i m sorry .
> who am i
Bot: you re not .
> what about weather in NY
Bot: there
> what's your hobby
Bot: i m sorry .
```

(b) RNN with Seq2Seq

```
> hey
Bot: hey
> what you doing
Bot: i m going to see you
> where are you from
Bot: gosh
> you want to come with me
Bot: no
> how much do you like me
Bot: i don t know
> who am i
Bot: gosh
> what about weather in NY
Bot: right here
> what's your hobby
Bot: gosh
> are you happy?
Bot: i m not sure
> Oh isee
Bot: right here
> bye
Bot: bye
```

(c) GRU

FIG. 6: COMPARISON OF ALL 3 ALGORITHMS

REFERENCES

[1] Peng H., Parikh A. P., Faruqui M., Dhingra B., and Das D. (2019). "Text generation with exemplar-based adaptive decoding," arXiv preprint arXiv:1904.04428.

[2] Luong M.-T., Pham H., and Manning C. D. (2015). "Effective approaches to attention-based neural machine translation," arXiv preprint arXiv:1508.04025.

[3] Sherstinsky A. (2020). "Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network," Physica D: Nonlinear Phenomena, vol. 404, p. 132306.

[4] Hochreiter S. and Schmidhuber J. (1997). "Long short-term memory," Neural computation, vol. 9, no. 8, pp. 1735–1780.

[5] Chung J., Gulcehre C., Cho K., and Bengio Y. (2014). "Empirical evaluation of gated recurrent neural networks on sequence modeling," arXiv preprint arXiv:1412.3555.

[6] Sutskever I., Vinyals O., and Le Q. V. (2014). "Sequence to sequence learning with neural networks," arXiv preprint arXiv:1409.3215.

[7] Cho K., Van Merri¨enboer B., Gulcehre C., Bahdanau D., Bougares F., Schwenk H., and Bengio Y. (2014). "Learning phrase representations using rnn encoder-decoder for statistical machine translation," arXiv preprint arXiv:1406.1078.

[8] Kingma D. P. and Ba J. (2014). "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980.

[9] Sojasingarayar A. (2020). "Seq2seq ai chatbot with attention mechanism," arXiv preprint arXiv:2006.02767.

[10] Graves A., Fern´andez S., and Schmidhuber J. (2005). "Bidirectional lstm networks for improved phoneme classification and recognition," in International conference on artificial neural networks. Springer, pp. 799–804.

[11] Sundermeyer Martin, Schl¨uter Ralf, and Ney Hermann (2012). "LSTM
Neural Networks for Language Modeling," In Interspeech, pages 194–197.

[12] Bengio Y., Ducharme Rejean, Vincent Pascal, and Janvin Christian (2000) . "A neural probabilistic language model," 13th International Conference on Neural Information Processing Systems, Pages 893–899.

[13] Bengio Y., Boulanger-Lewandowski N., and Pascanu R. (2013). "Advances in optimizing recurrent networks,", DOI: 10.1109/ICASSP.2013.6639349

[14] Csaky R. (2019). "Deep learning based chatbot models," arXiv reprint arXiv:1908.08835.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:05:21 UTC from IEEE Xplore. Restrictions apply.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/353659148

Analysis of Time Series Forecasting Techniques for Indian Automotive Industry

Research · July 2021

DOI: 10.13140/RG.2.2.12240.02564

CITATIONS
0

READS
24

2 authors, including:



Piyush Jain
Delhi Technological University
3 PUBLICATIONS **1** CITATION

SEE PROFILE

All content following this page was uploaded by Piyush Jain on 03 August 2021.

The user has requested enhancement of the downloaded file.

Analysis of Time Series Forecasting Techniques for Indian Automotive Industry

PIYUSH JAIN

Department of Production & Industrial Engineering, Delhi Technological University, Delhi, 110042

piyushjain584@gmail.com

PARV ARORA

Department of Production & Industrial Engineering, Delhi Technological University, Delhi, 110042

parvarora2000@gmail.com

Abstract: Automotive sector is one of the most crucial sectors for any developed or developing Economy. Unknown Demand for vehicles is a significant problem that decreases the productivity of the production environment. Thus, forecasting plays a vital role in planning the production and inventory of an industry. Time series forecasting is a type of forecasting in which last period data is used to determine the industry's trends and sales in the future. In this paper, the authors have deployed nine time series forecasting models for the Demand for vehicles in the Indian Automotive Sector. Authors have deployed Naïve Method, Simple Moving Average, Weighted Moving Average, Simple Linear Regression, Exponential Smoothening, Holt Winter linear Trend, Autoregression, Autoregression Moving Average, and Autoregression Integrated Moving Average models to compare the forecasted Demand and find the best model based on MAPE (Mean Absolute Percentage Error) value. This paper aims to scrutinize these nine models' ability to be implemented while performing market analysis or production planning of automobiles. The result of models helps analyze which model has greater accuracy and less percentage error, which will help production units.

Keywords: Demand Forecasting, Forecasting, ARIMA, Auto Regression, Moving Average, ARMA, Holt Winter

1. INTRODUCTION

The automotive sector is one of the developing sectors, where the most massive investments are made. It has an extensive business volume. Indian automobile production increased at 2.63% CAGR between FY16-FY20, with 26.36 million vehicles being manufactured in the country in FY20. Overall, domestic automobile sales increased at 1.29% CAGR between FY16-FY20, with 21.55 million vehicles being sold in FY20. It contributes significantly to the economic development of countries in the world. Today India’s automotive industry is expected to reach Rs 16 to 18 trillion by 2026. Therefore, it is crucial for companies in this sector to correctly manage their resources. To do that, companies should predict the future in the best possible way and anticipate possible issues. Automotive sectors are the critical sector of the Indian economy. There is a dire need for demand forecasting for production planning and predicting future trends in the industry. Demand forecasting plays a significant role in decision making for the upliftment of business. For the past several years, there has been immense development in techniques used for demand forecasting, and new and improved methods are being developed for more accurate results.

Forecasting is a method that uses past data and estimates the future growths. It is a very critical tool for informed decision making and planning. The goal of forecasting is to reduce the effect of random factors on the resulting situation as adequately as possible, as well as help the decision-maker choose the right strategy while making the decision. External factors such as taxes, incomes, etc. and internal factors such as quality, price, popularity of product play a vital role in determining the possible Demand in future. The basic steps in forecasting are determination of objectives of forecasting, selection of item to be forecasted, selection of suitable forecasting model, collection of data, validation of forecasting model and implementation of results.

Usually, when analyzing demand data that are considered time series and represent the historical market demand for the product in a given period of time, there is a necessity for forecasting – to determine the possible product demand in the future. Decision risks such as the minimum number of orders to be placed in the next sales cycle can be minimized using accurate forecasting. Traditionally, the process of the time series analysis is based on searching for patterns in a long period of time by analyzing the changes of values in moments of time; as a result, the future values of the analyzed object are extrapolated.

1.1. What is a Time Series?

Every sequential set of data-points constitutes a Time Series. Each Time Series has two main components: time and value assigned to the corresponding time step. Most common examples of time series are daily changes in stock prices of the company, heights of ocean tides, counts of sunspots, etc.

A time series that records the measurements of a single phenomenon/variable is called univariate, whereas the measurements of multiple phenomena/variables is called multivariate. Our paper focuses on forecasting and analyzing univariate time series of customer demand or number of registered vehicles every year in India, with the third largest road network in the world, the total number of vehicles in the fiscal year 2017 stood at 25.3 million

1.2 Types of Time Series

There exist many classification criteria for time series; in this part we will focus on time dependency and stationarity. Time dependency refers to the influence of past values on the newly observed values of the recorded variable/phenomenon, we can divide time series into two categories: long memory time series and short memory time series. Long memory time series have an auto-correlation function that decreases slowly, these time series describe a process that has slowly changed behavior. Long memory time series can usually be found in meteorological, geological data; an example of long memory time series is the evolution of the mean temperature on planets. Short term time series describe a process with a fast turnover and have an auto-correlation function that decreases rapidly as the longer we are from the present the less useful to the future the measure is. A typical example of these time series are financial data, like stock prices. For the second classification criteria: stationarity, we have stationary and non-stationary time series. Stationary time series are processes that have statistical properties (mean, variance) that do not depend on time. Time series that do not fit the aforementioned description are called non-stationary, these time series are ubiquitous in finance and the retail sector. These types of time series can be tough to predict without any type of pre-processing, that is why we use multiple techniques to stationaries these processes to be able to get better forecasts

1.3 Components of Time Series

There are different patterns of demand variation depending on the time series. These include linear variation (Fig 1), seasonal variation (Fig 2), cyclic variation (Fig 3), random variations (Fig 4).

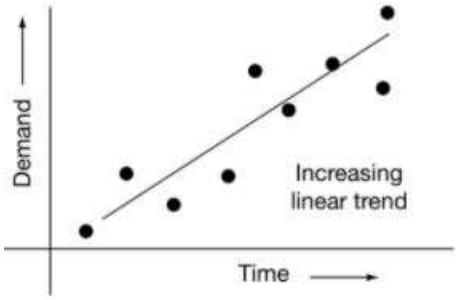


Figure 1

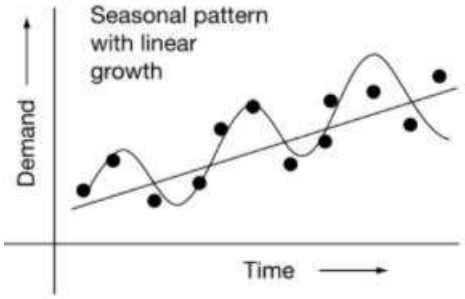


Figure 2

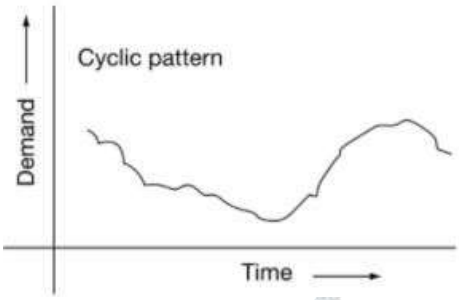


Figure 2

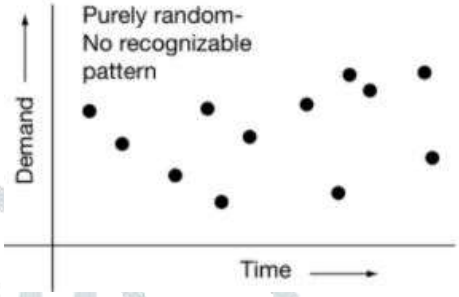


Figure 1

The components of the time series are as follows: -

- *Trends* are noted by an upward or downward sloping line.
- *Seasonality* is a data pattern that repeats itself over a period of certain duration. It leads to a rise or fall in Demand among the seasons. In a particular season demand may increase while in another season demand may decrease.
- *Cycle* is a data pattern that may cover several years before it repeats itself. Cyclic Nature in demand variations shows the same style in increasing or decreasing Demand in future.
- *Random* fluctuations result from random variations or unexplained causes. There are several reasons for random fluctuations in Demand, for example, anticipation of increasing the price or shortage of product in the near future increases the Demand.

2. LITERATURE REVIEW

Seher Arslankya in his research paper implemented time series analysis and artificial neural network methods to estimate sales for future months for the leading company in the automotive industry in Turkey. He examined the monthly data between January 2011 and July 2016 using multiple regression, moving average and artificial neural network model. Furthermore, he compared MAPE values for all models, which resulted in the ANN model giving the best result.

Arnis Kirshners did a comparative analysis of short time series processing methods intending to scrutinize these methods' ability to be used when analyzing short time series. The author has analyzed the moving average Method, exponential smoothening, and exponential smoothening with development trends resulting from moving average having the smallest squared error value but with large forecast smoothening for initial data.

Shamsul Masum elucidates on time series forecasting and its classification and approaches and strategies of time series forecasting. Furthermore, the author has demonstrated how an inappropriate point to point rolling forecast strategy leads to unrealistic outcomes and supports his argument with a comparative analysis of two

strategies using the ARIMA model. The author has concluded with the result of rolling single point outcome being deceptive for Euro Dollar Exchange Rates case considered.

Tamal Datta Chaudhuri proposed six different forecasting methods for predicting the time series index of the healthcare sector. The author has demonstrated a decomposition approach of time series for data from January 2010 to December 2016 and illustrated how the decomposition results provide us with useful insights into the behavior and properties exhibited by time series. The author observed that results from the ARIMA model with a horizon of 12 months came out to be the best model with the lowest RSME value, while the Holt Winters method with a horizon of 12 months has the highest RSME value.

Jaydip Sen, in his research paper, uses the Time series - decomposition based Method to analyze the past of the Indian realty sector and predict its future. He uses time series forecasting methods in R programming language to determine future results. He uses time series index value data of the Indian realty sector for six years from 2010-2016 month wise. The methods used for accurate predictions are: Holt Winters exponential smoothening and Autoregressive Integrated Moving Average. He analyses the results from the above-mentioned broad concepts and observes, which is the best one. With the result obtained, he argues that these can be immensely useful for portfolio managers and stock traders to buy or sell stocks at the correct time.

Samita sood elaborates how time series forecasting can be deployed in determining the future development of the Indian tourism industry from past secondary data. The author uses the data for the number of tourists in India from 1980 to 2020 and uses models such as ARIMA and Holt Winter to forecast foreign tourist travel and compare the two models' accuracy. On the basis of MAPE and RMSE, the authors conclude that Holt Winter is a more accurate model than ARIMA in this particular situation.

William R Huss implements univariate estimation techniques such as Holt winter exponential smoothing, Multiple regression, Linear regression to study the load on 49 largest electric utilities in the United States and forecast the load for future planning. The author uses electric utility data from 1972 to 1982, and the results indicate that for shorter periods, Holt Winters Exponential smoothing method is highly accurate, and for more extended periods, extrapolation of Linear regression horizons proves to be efficient.

Deepa, in this paper, reviews, and forecasts the Indian Motorcycle market using Time series forecasting. The author uses SARIMA (Seasonal autoregressive Integrated Moving Method) and Holt Winters Method for prediction. The author compares several years' data and uses MAE and MAPE method to determine which model is more accurate and concludes that both the models are pretty significant but Holt Winters method is numerically more precise than the SARIMA model. According to the author the studies can be further enhanced by using more such models.

3. METHODOLOGY

This paper aims to analyze various time series forecasting methods using a comprehensive data set of registered vehicles or vehicles purchased every year in India. Data for the past 15 years (2001-2017) was collected from the Ministry of Statistics and Programme Implementation, Government of India. Various time series forecasting such as naive Method, simple moving average, autoregression, autoregression integrated moving average, holt winter linear trend is deployed using Microsoft Excel to draw an analysis between the results obtained from forecast and actual values.

Abbreviations used in paper are as follows: -

D_t = Demand in time period t

F_t = Forecast for time period t

t = Time Period / Duration of 1 Cycle

n = number of the averaging period

W_t = weight applied to period t's demand

α = smoothing factor

$\bar{x}$ = mean value of x

Calculation of Mean Absolute Percentage Error (MAPE) :-

$$\text{MAPE} = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{D_t - F_t}{D_t}\right|$$

4. DIFFERENT TIME SERIES FORECASTING METHODS

4.1 Naive

Naive Method is the most basic forecasting method used in time series forecasting. In this Method the actual value of the last period is considered as a forecast of the current period without changing or implementing any statistical model. It is a special case of the moving average forecasting model where the number of periods used for smoothening is 1.

$$F_t = D_{t-1}$$

The forecast for the Tth period is equal to the Demand of the $T\text{-}1^{th}$ period.

4.2 Simple Moving Average

The simple moving average Method calculates the average value of actual Demand for some previous periods and forecast the Demand for the next period. This Method does not take into account the seasonality of data and it can only be used for forecasting the Demand of next period.

$$F_{t+1} = \frac{1}{n}\left(D_t + D_{t-1} + D_{t-2} + \ldots + D_{t+1-n}\right)$$

$$\Rightarrow F_{t+1} = \frac{1}{n}\sum_{i=\,t+1-n}^{i=t} D_i$$

Assumptions :

1. Observations older than *n* are not included at all.
2. All *n* past observations are treated equally.
3. It requires that *n* past observations be retained.

4.3 Weighted Moving Average

This Method is very similar to that of the simple moving average Method. The only main difference is that weights are allocated to different periods and the largest weight is provided to the most recent period and weights keep on decreasing as we move on to later periods.

$$F_{t+1} = W_1 D_t + W_2 D_{t-1} + W_3 D_{t-2} + \ldots\ldots + W_n D_{t+1-n}$$

Here $W_1 > W_2 > W_3 \ldots\ldots > W_n$ and the sum of all weights is 1 i.e. $W_1+W_2+W_3+\ldots W_n=1$.

Following assumptions are taken in WMA method:

1. Adjustments in the moving average to more closely reflect fluctuations in the data.
2. Weights are assigned to the most recent data.
3. Requires some trial and error to determine precise weights.

4.4 Exponential Smoothing

This is the most popular and efficient statistical Method. It is based on the principle that latest data should be weighed more heavily and 'smoothers' out cylindrical variations to forecast the trend. It assumes that as the data gets older it becomes less relevant. It is a more sophisticated WMA method that calculates the average of a time series by giving more weight to recent data. In this Method both the actual Demand and the forecast of previous Demand is used for calculation of forecasting.

$$F_{t+1} = \alpha D_t + (1 - \alpha) F_t = F_t + \alpha(D_t - F_t)$$

Here α is the smoothing factor and its value is decreased as we move towards the past. The decrease follows an exponential pattern thus the name Exponential Smoothing.

4.5 Simple Linear Regression

Simple Linear Regression is a mathematical technique that relates to one variable, that is, an independent variable , with another variable called the dependent in the form of a linear equation. The linear regression equation is -

$$y = a + b \cdot x$$

Where y is the dependent variable, a is the intercept , b is the slope of the line, and x is the dependent variable. The variables a and b are given by -

$$a = \bar{y} - b\,\bar{x}$$

$$b = \frac{\sum xy \; - n(\bar{y})\,(\bar{x})}{\sum x^2 \; - n\,(\bar{x})^2}$$

Where a is constant and b of the coefficient of variable x , x is time and y is Demand.

The coefficient of correlation shows the strength of correlation between 2 variables. The linear regression model of forecasting is only used when coefficient of correlation(r) is high i.e. near to 1. Correlation variable is a measure of strength of the relationship between the independent (time) and dependent (Demand) variable. It is given by :

$$r = \frac{\sum_{t=1}^{n}(x_t - \bar{x})\,(y_t - \bar{y}\,)}{\sqrt{\sum_{t=1}^{n}(x_t - \bar{x})^2}\sqrt{\sum_{t=1}^{n}(y_t - \bar{y})^2}}$$

4.6 Auto Regression

An auto regressive model predicts future behavior based on past behavior. It is used for forecasting when there is a sum correlation between values in a time series and values that precede and succeed them. The process is a linear regression of the current series's data against one or more past values in the same series. An AR(p) model is an auto regressive model where specific lagged values of y_t are used as predictor variables. The value for p is

called the order. The order of an autoregression is the number of immediately preceding values in the series that are used to predict the value at the present time. The general equation for AR(p) model is as follows:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \ldots\ldots + \phi_p y_{t-p} + \varepsilon_t$$

$$\phi_0 = (1 - \sum_{i=1}^{p} \phi_i)\, \mu$$

Where μ refers to process mean and ε_t refers to the white noise or shock / randomness in the data which ranges from (0 , σ).

AR(1) model refers to an autoregressive process in which the current value is based on the immediate preceding value. Equation for first order auto regression model is as follows :-

$$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$$

4.7 Auto Regression Moving Average (ARMA)

In the statistical analysis of time series, autoregressive–moving-average (ARMA) models provide a parsimonious description of a (weakly) stationary stochastic process in terms of two polynomials, one for the autoregression (AR) and the second for the moving average (MA). It is the basic model for analyzing a stationary time series. A stationarity time series is defined formally in terms of the behavior of the autocorrelation function (ACF) through Wold's decomposition. Several simple cases of the ARMA model are then introduced and analyzed, with the partial autocorrelation function (PACF) also being defined, before the general model is introduced.

Finding appropriate values of p and q in the ARMA(p,q) model can be facilitated by plotting the partial autocorrelation functions for an estimate of p, and likewise using the autocorrelation functions for an estimate of q. Further information can be gleaned by considering the same functions for the residuals of a model fitted with an initial selection of p and q.

$$y_t = \phi_0 + \varepsilon_t + \sum_{i=1}^{p} \phi_i\, y_{t-i} + \sum_{i=1}^{q} \theta_i\, \varepsilon_{t-i}$$

ARMA(1,1) model refers to an autoregressive moving average process in which the current value is based on the immediate preceding forecast and residual value. The equation for the first-order auto regression model is as follows:-

$$y_t = \phi_0 + \varepsilon_t + \phi_1\, y_{t-1} + \theta_1\, \varepsilon_{t-1}$$

4.8 Auto Regression Integrated Moving Average (ARIMA)

An autoregressive integrated moving average model is a form of regression analysis that gauges the strength of one dependent variable relative to other changing variables.

Each component functions as a parameter with a standard notation. For ARIMA models, a standard notation would be ARIMA with p, d, and q, where integer values substitute for the parameters to indicate the type of ARIMA model used. The parameters can be defined as:

p: the number of lag observations in the model; also known as the lag order.

d: the number of times that the raw observations are differenced; also known as the degree of differencing.

q: the size of the moving average window; also known as the order of the moving average.

Differencing (back- shift) operator :

$$\Delta y_t = y_{t-1}$$
$$(1-\Delta)\, y_t = y_t - y_{t-1}$$
$$\Delta^d y_t = \Delta^{d-1} y_{t-1}$$
$$= y_{t-d}$$

Re-writing AR(p), MA(q) and ARMA(p,q) using the back-shift operator will simplify the mathematical representation of the ARIMA model.

4.9 Holt Winter Linear Trend model

Holt (1957) extended simple exponential smoothing to allow the forecasting of data with a trend. This Method involves a forecast equation and two smoothing equations (one for the level and one for the trend):

Forecast equation $y_{t+h|t} = \ell_t + hb_t$

Level equation $\ell_t = \alpha y_t + (1-\alpha)(\ell_{t-1}+b_{t-1})$

Trend equation $b_t = \beta(\ell_t - \ell_{t-1}) + (1-\beta)b_{t-1},$

where ℓ_t denotes an estimate of the level of the series at time t, b_t denotes an estimate of the trend (slope) of the time series at time t, α is the smoothing parameter for the level, $0 \le \alpha \le 1$, and β is the smoothing parameter for the trend, $0 \le \beta \le 1$.

As with simple exponential smooth (SES), the level equation here shows that ℓ_t is a weighted average of observation y_t and the one-step-ahead training forecast for time t, given by $\ell_{t-1}+b_{t-1}$. The trend equation shows that b_t is the weighted average of the estimated trend at time t based on $\ell_t - \ell_{t-1}$ and b_{t-1}, the previous estimate of the trend.

The forecast function is no longer flat but trending. The h-step-ahead forecast is equal to the last estimated level plus h times the last estimated trend value. Hence the forecasts are a linear function of h.

5. NUMERICAL ILLUSTRATIONS

Year	Actual Sales	Volume Produced	Naïve Method	Simple Moving Average	Weighted Moving Average	Exponential Smoothening	Holt Winter Linear Trend	Simple Regression	Auto Regression	ARMA	ARIMA
2010	12,795	17,916	9,598	8,444	8,555	9,198	10,887	7,332	10,095	10,858	9,370
2011	14,120	20,382	12,795	10,346	9,842	11,715	12,556	7,552	11,030	11,856	11,935
2012	17,625	20,626	14,120	12,171	11,360	13,398	15,026	7,771	11,732	12,946	12,279
2013	18,422	21,500	17,625	14,847	13,681	16,355	17,130	7,990	12,717	14,136	13,918
2014	19,724	23,358	18,422	16,722	15,636	17,801	19,062	8,210	13,602	15,436	15,861
2015	20,469	24,016	19,724	18,590	17,445	19,147	20,628	8,429	14,687	16,856	17,182
2016	21,863	25,314	20,469	19,538	18,812	20,072	22,118	8,648	15,752	18,405	19,366

2017	24,981	29,094	21,863	20,685	20,147	21,325	24,169	8,867	16,974	20,098	21,477
2018	26,266	30,914	24,981	22,438	21,855	23,883	26,019	9,087	18,231	21,946	23,547
2019	21,546	26,362	26,266	24,370	23,540	25,550	25,530	9,306	19,626	23,963	26,148

Table 1 : Actual Sales , Vehicles Produced and Forecasting Result for Different Time Series Forecasting Techniques (all values are in thousands)

In Table 1 values of forecasted results for different techniques have been shown along with the actual sales and actual vehicle produced for each year. For simple moving average Method, calculations were conducted to find the most optimal period duration and the 3-period moving average gave the least MAPE value. For the exponential smoothening Method, the most optimal solution is for the smoothening factor 0.69. Analysis for Holt Winter Linear Trend gave minimum MAPE value for alpha :0.35 and beta:0.25. While considering the case for simple regression , the intercept value came out to be -1658 and x variable value 219. A comparison was drawn between different AR order models [AR(1) , AR(2) , AR(3)]. AR(2) model was the most precise model and had least MAPE value. Comparing ARMA [(1,1) ; (1,2) ; (2,2)] order models for same data set, ARMA(1,1) had least MAPE value. Forecasting using different ARIMA order models without constant [(1,2,2) ; (1,1,1) ; (2,2,2) ; (1,2,3) ; (3,2,1) ; (3,2,3) ; (3,3,3) ; (4,3,2)] were quite comparable but ARIMA(3,3,3) produced most precise result among all. All the models are trained using data for the past 40 years (1970 – 2009) and were deployed to forecast values for the next 10 years (2010 – 2019). Table 2 highlights MAPE value for all the techniques used for forecasting with the most optimum order or smoothening factor among themselves.

Forecasting Technique	MAPE
Naïve Method	11%
Simple Moving Average	19%
Weighted Moving Average	22%
Exponential Smoothening	15%
Holt Winter Linear Trend	8%
Simple Regression	57%
Auto Regression	27%
ARMA	18%
ARIMA	19%

Table 2 : MAPE for Different Time Series Forecasting Techniques

6. RESULT AND CONCLUSIONS

According to the conducted experiments the Holt Winters Linear Trend achieves the minimum Mean Absolute Percentage Error (MAPE) and Simple Regression has the maximum MAPE value for the chosen data set of vehicles sold in India for the past 50 years (1970 – 2019). The study demonstrates that Holt Winter Method is the most precise and accurate model for Indian Automotive industry in all available time series forecasting techniques. The Holt-Winters model's relative ease of use makes the model useful in forecasting comprehensive market trends. Some results obtained from other time series forecasting techniques were quite comparable. Methods such as Exponential Smoothening , ARMA , Moving Average , ARIMA also had relatively less MAPE value but Holt Winters proved to be the best among them.

The actual vehicles sold in an year were compared with the actual production of the number of vehicles produced that year. The average overproduction percentage came out to be a staggering 18%. The analyzed results emphasize that there is a necessity for better and improved production planning strategies to minimize the losses faced by companies due to over or under production. The primary step involved in the improvement of production strategy is to have precise forecasted results. Improving on the accuracy of estimation is imperative to the reasonable allocation of scarce resources, this signifies the call for enhanced forecasting techniques and the research will help Indian automobile industry and parts manufacturers to build effective strategies.

REFERENCES

[1]. Sen J., Chaudhary T.D. (2017) , A Time Series Analysis-Based Forecasting Framework for the Indian Healthcare Sector, Journal of Insurance and Financial Management, 3(1)

[2]. Sood S, Jain K (2017) ,Comparative Analysis of Techniques for Forecasting Tourists' Arrival , Journal of Tourism & Hospitality p.p. 285 , 6(3) doi: 10.4172/2167- 0269.1000285

[3]. **S**. Masum, Y. Liu and J. Chiverton, "Comparative analysis of the outcomes of differing time series forecasting strategies," In: proceedings 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Guilin (2017), pp. 1964-1968, doi: 10.1109/FSKD.2017.8393069.

[4]. Baccar Y.B. (2019), Comparative Study on Time Series Forecasting Model, University of Paris-Saclay, Master Thesis

[5]. Trading Economics Homepage, https://tradingeconomics.com/india/car-registrations , Last accessed on 2020/10/18

[6]. Puthran D., Kumar K.S.K., Muniyappa M., Prasad H.C. (2014), Comparing SARIMA and Holt-Winters' forecasting accuracy with respect to Indian motorcycle industry, Transactions on Engineering and Sciences, 5(2)

[7]. Kirshners A., Borisov A. (2012), A Comparative Analysis of Short Time Series Processing Methods, Information Technology and Management Science, 15

[8]. Huss, W. R. (1985) , Comparative analysis of company forecasts and advanced time series techniques using annual electric utility energy sales data , International Journal of Forecasting, p.p. 217–239 , 1(3) , doi:10.1016/0169-2070(85)90004-4

[9]. Zhang, X., Liu, Y., Yang, M., Zhang, T., Young, A. A., & Li, X. (2013), Comparative Study of Four Time Series Methods in Forecasting Typhoid Fever Incidence in China, PLoS ONE, 8(5)

[10]. Jha K., Sinha N., Sarkar A.K., Arkatkar S.S. (2013), Modeling Growth Trend and Forecasting Techniques for Vehicular Population in India, International Journal for Traffic and Transport Engineering, 3(2)

[11]. Arslankaya S., Öz V. (2018), Time Series Analysis on Sales Quantity in an Automotive Company and Estimation by Artificial Neural Networks, SAKARYA UNIVERSITY JOURNAL OF SCIENCE, p.p. 1482-1492, 22(5)

[12]. Christoph K., Pircher M., Sharma S. (2004), Univariate Time Series Forecasting, University Of Vienna

[13]. Sen J. (2017), A Time Series Analysis - Based Forecasting Approach for the Indian Realty Sector, International Journal of Applied Economic Studies, 5(4)

[14]. Ministry of Statistics and Programme Implementation Homepage, http://mospi.nic.in/statistical-year-book-india/2018/189, Last accessed on 2020/11/10

View publication stats

Fibers and Polymers
DOI 10.1007/s12221-021-3178-9

ISSN 1229-9197 (print version)
ISSN 1875-0052 (electronic version)

Antimicrobial Finishing of Cotton Textiles using Silver Intercalated Clay

Rosina Begam[1,2], Mangala Joshi[1], and Roli Purwar[1,3]*

[1]*Department of Textile Technology, Indian Institute of Technology, Hauz Khas, New Delhi-110016, India*
[2]*Department of Chemistry, B.S. Abdur Rahman Crescent Institute of Science and Technology, Vandalur, Chennai-600 048, Tamil Nadu, India*
[3]*Discipline of Polymer Science and Chemical Technology, Department of Applied Chemistry, Delhi Technological University, Delhi-110042, India*

(Received February 22, 2021; Revised April 20, 2021; Accepted May 11, 2021)

Abstract: This paper describes the potential use of silver intercalated clay as an antimicrobial finishing agent for cotton textiles meant for health and hygiene applications. Montmorillonite clay (2:1) was intercalated with silver ion through ion exchange reaction. The modified clay was characterized using wide angle X-ray diffractometer (WAXD), Energy dispersive X-ray spectroscopy (EDX) and inductive coupled plasma spectroscopy (ICP). Results of WAXD revealed that the clay was partly exfoliated after ion exchange reaction. EDX and ICP analysis confirmed the presence of silver ion in modified clay. The silver intercalated clay was integrated on the cotton fabric using pad dry cure method along with low formaldehyde dimethyloldihydroxy ethylene urea (DMDHEU). The antimicrobial and mechanical properties of the finished fabric were studied. The fabric showed excellent antimicrobial activity up to 10 machine washes. The silver ion releases from the fabric and inhibited the growth of bacteria on the surface of the fabric. The release of silver ion from the finished fabric at different perspiration and saline water were analyzed. It was found that the silver ion releases faster in alkaline perspiration medium compared to acid perspiration medium.

Keywords: Montmorillonite, Silver intercalated clay, Antimicrobial finish, DMDHEU, Alkaline perspiration

Introduction

Antimicrobial textiles have created a special attention all over the world in the last decade due to the awareness towards sanitation, personal protection from the spread of bacteria, subsequently, the infection and unpleasant odor. A large number of commercially available organic and inorganic antimicrobial agents are being used by the industry for the protection purpose [1].

Silver and its compounds have found potential use in textiles and most extensively studied due to its wide spread antimicrobial properties, biocompatibility and non toxicity to human cells at optimum concentration which is effective against microorganism. Silver deserves special attention due to its unique property that the bacteria are not able to develop their resistance to silver as in case of antibiotics [2]. The mechanism for the antimicrobial action of silver ions is not yet properly understood. However, it is suggested that the silver ions penetrate inside the bacteria cell, reacts with DNA molecules which turns it into condensed form and loses its ability for replication, leading to cell death. It was also reported that the heavy metals react with protein by getting attached with the thiol groups and the protein get inactivated [3].

Silver is used on cotton textiles mostly in the form of silver salt [4], silver nanoparticles [5-7] and silver complexes [8]. Silver and its various forms have been applied on to the cotton textiles using exhaust, pad dry cure and coating methods. The attempts were made to apply silver on the cotton textiles using sol gel [9], sono-chemical coating [10,11] and chemical plating [12] and DC electrical discharges between two electrodes techniques [13]. The antimicrobial activity of silver based agents on textiles is mostly attributed to their gradual and persistent release from textiles in to the surrounding, where they act as a cidal to a wide spectrum of microorganisms.

The organic modified nano clay as a nano filler finds use in the preparation of hybrid polymer nano composite fibre due to improved mechanical, thermal, electrical, barrier, optical and dyeing properties [14]. The organic modified clay is generally prepared through ion exchange of the metal cations within the clay with organophilic cations such as quaternary ammonium salts. It has been reported recently that the organoclay containing cationic surfactant or quaternary ammonium salt show excellent antibacterial properties [15]. The use of montmorillonite clay as a supported medium or carrier for antimicrobial agents such as silver, copper, quaternary ammonium compounds, chitosan etc. was studied by various researchers in recent years, due to their potential use in water treatment, food packaging, and drug delivery material [16-18]. In our previous study we have developed bioactive sutures by incorporating silver and copper intercalated clay in melt spun filaments [19]. The use of silver intercalated clay has great potential as an antimicrobial finishing agent for textiles. To best of our knowledge, there is no study on the application of silver intercalated clay as finishing agent on textiles is reported. The major advantage of silver intercalated clay finished fabric over silver nano particles finished fabrics that the silver ion release is controlled and sustainable manner which

*Corresponding author: roli.purwar@dtu.ac.in

© The Korean Fiber Society for Fibers and Polymers and Springer 2021

enhance its durability. In this study, montmorillonite clay (2:1) was modified with silver ion through ion exchange reaction and characterized. The silver modified biocidal clay was integrated on cotton textile using pad dry cure method along with low formaldehyde dimethylol dihydroxy ethylene urea (DMDHEU). The antimicrobial and mechanical properties of the finished fabric were studied in detail.

Experimental

Materials

A desized and bleached cotton fabric with mass per unit area of 125 g/m^2 was used in this study. Sodium Montmorillonite clay was procured from Kunimine Industries Co. Ltd., Japan. Low formaldehyde based DMDHEU Crosslinking agent (Arkofix) was procured from, Clariant India Ltd., India. Reagent grade chemicals namely, Silver Nitrate (Merck), Acetone (Merck), Magnesium Chloride - (Merck), Nutrient Broth (Hi Media), Agar Agar (Qualigen) and deionized water were used.

Methods

Preparation of Silver Modified Montmorillonite Clay

The silver intercalated clay was prepared according to the method described by [18]. In brief, the ball milled montmorillonite clay was treated with silver nitrate solution at room temperature (27 °C±3 °C) for 7 days under constant stirring condition at 500 rpm. The concentration of silver nitrate was used as six times (6 CEC) of the cation exchange capacity of clay. The possible ion exchange reaction takes place between cations namely, Na^+, K^+ and Ca^{++} present within the clay and the Ag^+ present in silver nitrate solution as shown in Scheme 1.

Characterization of Clays

The change of d-spacing for intercalated clay was studied on wide angle X-ray diffractometer (Philips) using a Ni-filtered CuKα radiation of 1.5418 Å. A ZEISS (Model: EVO 50) SEM equipped with EDX attachment was used for the analysis of quantitative elemental composition of the clays. The paste of different clays were made in isopropanol and mounted on EDX sample holder. The samples were stubbed and coated with carbon to prevent charging by the electron beam. Inductively coupled plasma-optical emission spectrometry (ICP-OES Perkin Elmer OPTIMA 3000) was used to measure the quantity of silver intercalated on modified clay samples. Clay samples of 0.5 g was incinerated in a digitally controlled furnace. Temperature was gradually increased to 600 °C and then maintained for 60 minutes. Remaining ashes were dissolved in concentrated nitric acid in the 50 m*l* volumetric flask, which was then made up to 50 m*l* using double-distilled water. All solutions were preserved in plastic containers at room temperatures of 27 °C±2 °C. Measurements for each sample were performed in triplicate and average results are reported.

$$Na_i.M_{x+i}(Al_2\text{-}Mg_x)\ Si_4\ O_{10}\ (OH)_2\ .nH_2O + (n+x)AgNO_3$$

(Sodium Montmorillonite) (Silver Nitrate)

$$\Downarrow \text{ 500 rpm, 30° C, 7 days}$$

$$nAg\ M_x\ (Al_2\text{-}Mg_x)\ Si_4\ O_{10}\ (OH)_2.\ nH_2O + Na_i\ M_i.\ xAg\ NO_3$$

(Silver Montmorillonite)

Scheme 1. Ion exchange reaction of sodium montmorillonite and silver nitrate.

Integration of Silver Intercalated Clay on Cotton Textiles

The silver intercalated clay was applied on the cotton textiles using pad dry cure method along with low formaldehyde DMDHEU cross linking agent. Scoured and bleached cotton fabric samples were dipped in a nano clay dispersion containing biocidal clay, 5 g/*l* of crosslinking agent (Arkofix, Clariant), catalyst $MgCl_2$ (4 % on the weight of cross linking agent) for 2 minutes and then padded through padding mangle under constant pressure to achieve the percent expressions of 70 %. The samples were dried at 85 °C for 5 minutes and cured at 140 °C for 2 minutes. After finishing, the fabric samples were conditioned at 27 °C± 2 °C, 65 %±3 % relative humidity for 24 hours. The weight add of the finished samples was determined using following formula.

$$\text{Weight add on \%} = (W_2 - W_1) / W_1 \times 100$$

W_1 is the weight of the unfinished fabric (g)
W_2 is the weight of the finished fabric (g)

Characterization of Finished Fabric

The infrared spectra of the fabric samples were recorded between 400-4500 cm^{-1} on Perkin-Elmer Spectrum-BX FTIR system (Massachusetts, USA) using Potassium Bromide KBr pellet technique. The KBr pellets were prepared by grinding one part of the sample with nine parts of spectral grade KBr and pressed in an evacuated die under suitable pressure to get pellets. The ZEISS (Model: EVO 50) SEM equipped with EDX attachment was used for the qualitative elemental composition of the biocidal clay finished fabric. The fabric samples were mounted on EDX sample holder and coated with carbon to prevent charging in the electron beam.

Evaluation of Antimicrobial Activity of Finished Fabrics

Antibacterial activity of the fabrics was qualitatively studied against Gram positive bacteria *Staphylococcus aureus* and Gram negative bacteria *Escherichia coli* by disc diffusion test (AATCC-90). The bacteria strain was cultured

on Luria broth solution at 37 °C for 24 h. The stock solution consists of 2-8×10^7 CFU/m*l*. An amount of 10 µ*l* (diluted to 10^{-5} CFU/m*l*) of inoculum was spread on the Luria agar plate. The sterilized fabric disc (24 mm) was placed over the surface of the inoculated agar plates. The sodium chloride solution (20 µ*l* 0.1 % w/v) was added to the agar medium to act as counter ions for diffusion of Ag^+ ions. The zone of inhibition was measured after 24 h of incubation at 37 °C.

The finished fabric samples were subjected to modified colony counting method. The brief of the test procedure is given below. Initially the fabric sample testing was conducted as disc diffusion test. After 24 hours of incubation at 37 °C the sample was taken out and placed in conical flask containing 10 m*l* of distilled water. The flask was shaken for 1hour at 200 rpm on a laboratory shaker to release the bacteria from the fabric. Serial dilution of the liquid was made in sterilized distilled water. Dilution of 10^{-4} and 10^{-5} were used for colony counting. 10 µ*l* of inoculum were spread on to the agar plate and plates were incubated at 37 °C for 24 hours. After incubation, bacterial colonies were counted.

$$\text{Antibacterial activity (\%)} = (A - B)/A \times 100$$

where A: Number of colonies in control sample
B: Number of colonies in treated samples

Wash Durability

The antibacterial activity of the fabric with subsequent laundering was analyzed. The washing of the fabric was carried out at 50 °C for 45 minutes in launder-o-meter with 5 g/*l* LISSAPOL N.

Silver Ion Diffusion Studies from the Finished Fabrics

The silver ion diffusion from the finished fabric was studied in two different kind of simulated perspiration solution and saline water. The simulated perspiration solutions namely, acidic (pH 5.5) and alkaline (pH 8) were prepared by the method described in the "Test method for color fastness to perspiration" [20]. For the preparation of acidic perspiration solution, L-histidine hydrochloride monohydtrate (0.5 g), sodium chloride (5 g) and sodium dihydrogen orthophosphate anhydrous (NaH_2PO_4) (2.2 g) were dissolved in distilled water, mixed with sodium hydroxide 0.1 M to adjust the pH to 5.5. For the preparation of alkaline perspiration solution, L histidine hydrochloride monohydrate (0.5 g), sodium chloride (5 g), Disodium hydrogen phosphate dehydrate ($Na_2HPO_4 \cdot 2H_2O$) (2.5 g) were dissolved in distilled water and mixed with sodium hydroxide 0.1M to adjust the pH to 8.0. For silver ion elution from the finished fabric, 1 gram of fabric sample was placed in 20 m*l* of perspiration/saline solution in Erlenmeyer flask at 200 rpm using incubated shaker (Biogenisis) at 30 °C for silver ion release. After specified time, the solutions were taken out, filtered through Whatman filter paper. The release of silver ions from the fabrics was simulated to perspiration/saline solution and the concentration of released silver was determined by using ICP-ES instrument (JOBIN YVON JY 24 Sequential ICP-ES). The test was performed with 1 % silver intercalated clay finished fabric. The amount of silver present on the fabric sample was evaluated by ICP and found to be 320 ppm. The finished fabric samples were inserted in the different perspiration medium as well as saline water, the release of silver ion from the samples were evaluated after 1, 4, 8, 12, 24 hours using ICP.

Performance Properties of Finished Fabric

Samples were conditioned (27 °C±2 °C and 65 %±3 % relative humidity for 24 hours) prior to study of mechanical properties. The tensile strength of the finished fabric was evaluated according to ASTM D 5035 method. Standard AATCC-66 test method was used to determine the crease recovery angle (CRA). For each sample, 10 specimens were tested and the average value is reported. Bending length measurement was carried out in both warp and weft direction of the treated and untreated samples using Eureka bending length tester according to ASTM D1388-08. The color on the finished fabric samples was measured by Gretag Macbeth, Color Eye-7000-A using Color-eye Control software.

Results and Discussion

Characterization of Clay

The average crystal structures of sodium montmorillonite and silver intercalated clays were analyzed by X-ray diffraction. Figure 1 shows the X-ray diffractogram of sodium montmorillonite (Na-MMT) and silver modified montmorillonite (Ag-MMT). The X-ray diffractogram of Na-MMT showed the characteristic d spacing of 12.49 Å at 2θ value of 7.12 °. Small quantities of quartz, at 26.68 ° (2θ), and anorthite, at 20.07 °, 26.71 ° and 28.63 ° (2θ), were also observed. It has been observed that the peak of 011 plane became less intense in XRD diffractograms of silver modified montmorillonite clay. Decrease of the basal diffraction intensity indicates the partly exfoliated layered structure. Similar studies have been reported by Praus *et al.* [18], with increase in concentration of silver nitrate solution,

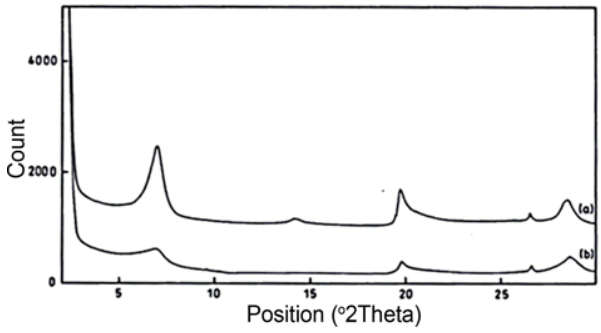


Figure 1. X-ray diffraction of (a) sodium montmorillonite and (b) silver modified montmorillonite.

Table 1. Elemental chemical analysis of clays by EDX

Elements	Si	Al	O	Mg	Fe	Mn	Na	K	Ca	Ag
Na-MMT	22.71	9.55	64.78	1.48	2.53	0.61	2.01	0.04	0.55	0
Ag-MMT	25.5	6.71	58.19	1.47	2.84	0.48	0	0	0.03	10.53

the content of Ag^+ in MMT increases and the layered MMT structure gradually collapsed and exfoliated.

Elemental chemical analysis of sodium montmorillonite, and silver intercalated clays were carried out using EDX and the results are summarized in Table 1. Na-MMT clay contains Na^+, K^+, Ca^+, ions and the major ion is Na^+ accounting for about 2 % of the clay weight. After the ion exchange reaction these interlayer cations are replaced by Ag^+ ions. The amount of silver ion present on silver intercalated clay is about 10.5 wt %. The silver intercalated clay samples showed a decrement in Al^{+++}, Na^+, K^+ and Ca^{++} content after ion exchange reaction. Therefore, these are the preferential exchangeable ions for Ag^+. Magaña *et al.* have prepared various silver loaded montmorillonite clays with maximum 8.37 wt% [16], whereas in this work, higher loading of silver about 10.5 wt% is achieved.

Characterization of Silver Intercalated Clay Finished Fabric

The attachment of the silver modified montmorillonite on the cotton fabric surface using DMDHEU cross linking agent was analyzed using FTIR technique. The FTIR spectra of the fabrics were taken after first wash of the samples to avoid the superficial attached DMDHEU as well as silver montmorillonite. Figure 2 shows the FTIR spectra of untreated, DMDHEU treated and silver montmorillonite treated cotton fabrics. There was no significant observation obtained from the three different FTIR spectra. This may be due to the higher content of cellulose which masks the peak of additional groups. There was slight change in finger print

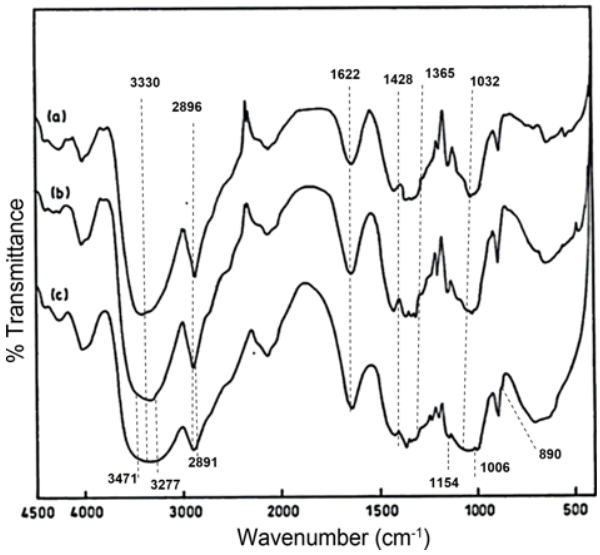


Figure 2. FTIR Spectra of the fabrics (a) untreated, (b) DMDHEU finished fabric, and (c) silver montmorillonite with DMDHEU treated fabric.

Table 2. FTIR peaks and their interpretations

DMDHEU treated-untreated cotton fabric (cm^{-1})	Silver montmorillonite treated-untreated cotton fabric (cm^{-1})	IR peak interpretation
3471	3471	H_2O-H_2O-H & Strong hydroxyl band
3277	3277	Hydroxyl band
2891	2891	C-H stretching associated with DMDHEU
-	1154	Si-O-Si
-	1006	Si-O
-	890	Al-O-H

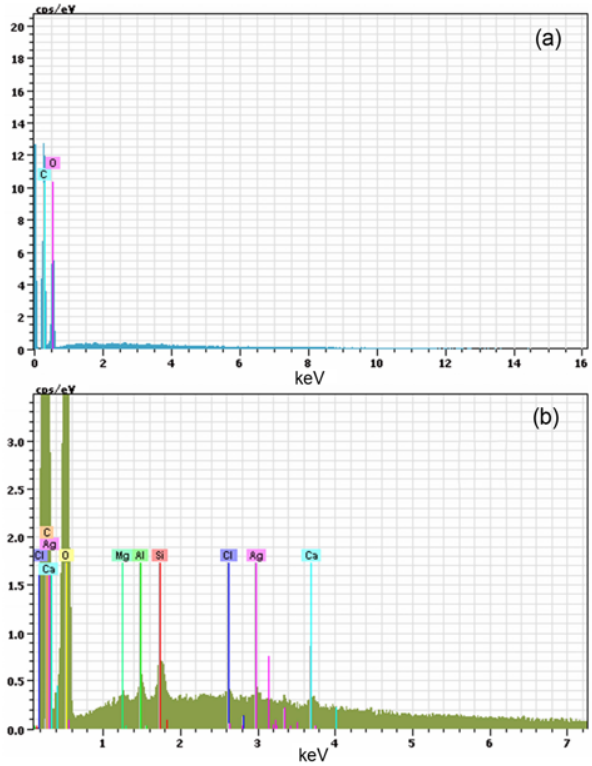
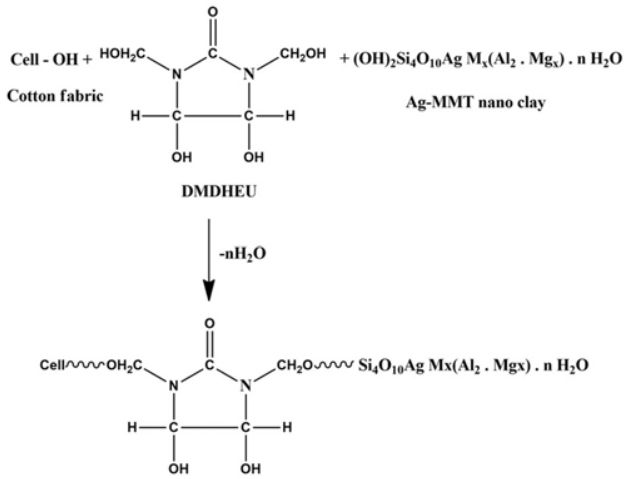


Figure 3. EDX spectra (a) untreated fabric and (b) silver montmorillonite finished fabric.



Scheme 2. Mechanism of cross-linking of cotton with Ag MMT nanoclay.

region observed. Therefore, the IR spectrum of untreated cotton fabric was subtracted from the IR spectra of DMDHEU treated fabric and Silver montmorillonite treated fabric. major peak observed is tabulated in Table 2.

The additional peaks in silver montmorillonite treated fabrics have been observed in the subtraction spectra, the additional peaks are associated with the Si-O-Si bond (1154 cm^{-1} and 1006 cm^{-1}) as well as Al-O-H (890 cm^{-1}). This indicates that the silver montmorillonite is chemically attached to the cotton textiles using DMDHEU cross linking agent.

The EDX spectra of silver intercalated clay treated and untreated fabrics are shown in Figure 3. The EDX spectra of untreated fabric *i.e.* cellulose shows only C and O peaks. The EDX spectra of the silver intercalated clay treated fabrics show the presence of silver ion on the surface of the cotton fabric (Figure 3). This also indicates that the silver intercalated clay chemically attached to the cotton fabric. The possible mechanism of attachment of silver intercalated clay on cotton fabric is shown in Scheme 2.

Antibacterial Activity of the Silver Intercalated Clay Finished Fabric

The antimicrobial activity of the untreated and silver intercalated clay treated fabric samples were qualitatively and quantitatively evaluated using disc diffusion and modified colony counting method respectively against *Staphylococcus aureus* and *Escherichia coli*. The antibacterial testing of the samples was conducted after 1st wash. The results of disc diffusion test are shown in Figure 4. The untreated cotton fabric showed huge growth of bacteria on the surface of the fabric. The silver intercalated clay treated fabric samples showed a clear zone of inhibition around the fabric. The silver ion diffused out from the fabric, killed the bacteria around the fabric and created a zone of inhibition.

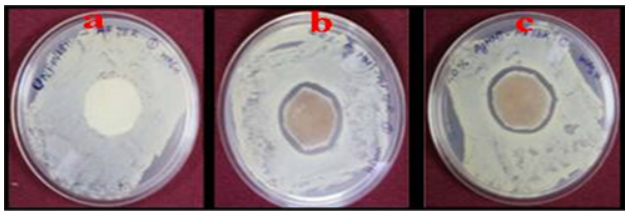


Figure 4. Antibacterial activity of 1.0 % AgMMT treated fabrics after first wash by disc diffusion method (a) untreated, (b) *S. aureus*, and (c) *E. coli.*

The zone of inhibition of the samples increases with the increase in concentration of the silver intercalated clay. The zones of inhibition created by different samples are summarized in Table 3. The wash durability of the biocidal clay finished fabrics with respect to antibacterial property was evaluated. It was found that the all the fabric samples concentration ranging from 1 % to 2 % created zone of inhibition even after 5 washes. After 10th wash, the samples were not showing distinct zone of inhibition however there was no visible bacterial growth on the surface of the fabric. To check any growth on the surface of the fabric after 10th wash, the samples were tested using modified agar plate method as mention in methods section. The results are summarized in Table 4. The untreated fabric showed the bacterial growth in the range of 1-2×10^8 CFU/m*l*. The fabric treated with 0.25 % silver intercalated clay did not show a distinct antibacterial activity as compared to the control fabric. But the 0.5, 1.0 and 2.0 % silver intercalated clay finished fabrics showed excellent bacterial reduction, after

Table 3. Antibacterial activity of fabrics against *Staphylococcus aureus* and *Escherichia coli*

Sample	Concentration of biocidal clay (%)	Zone of inhibition[#] (mm)
Untreated	-	24
S. aureus	1.0	30
E. coli	1.0	30

[#] Initial disc size diameter: 24 mm.

Table 4. Washing durability of fabrics after 10th wash by colony counting method

Sample	Concentration of biocidal clay (%)	*Staphylococcus aureus* (CFU/m*l*)	*Escherichia coli* (CFU/m*l*)
Untreated	-	1-2×10^8	1-2×10^8
Sample 1	0.25	1-2×10^8	1-2×10^8
Sample 2	0.50	1-2×10^7	1-2×10^7
Sample 3	1.0	1-2×10^6	1-2×10^6
Sample 4	2.0	1-2×10^3	1-2×10^3

10th washing cycle, which is equivalent to 25 home laundering AATCC Test Method 61-1996. So far maximum wash durability with respect to antibacterial activity achieved by various researchers for silver ion and silver nano particle finished cotton fabric is about 80-90 % after 10 machine wash [21,22]. In our study we have achieved 100 % antibacterial activity after 10 machine wash with only 0.5 wt% silver intercalated clay. This shows that the silver ion releases from the clay galleries in controlled and sustainable manner.

Silver Ion Release from the Silver Intercalated Clay Finished Fabric

The release of silver ion from the finished fabric in different perspiration medium (acidic and alkaline) as well as in water was evaluated using by ICP. Figure 5 shows the release of silver with respect to time in different medium. There is continuous release of silver ion with respect to time. The initial amount of silver ion releases higher in alkaline medium (54 ppm) due to the presence of more sodium ions present in alkaline medium compared to others which can interchange to silver ions more as compared to that of acidic (15 ppm) and saline medium (10 ppm). After initial release, the rate of silver ion release from the fabric is very slow in acidic and saline medium. However the rate of silver ion release is higher in the alkaline medium. This may be due to the alkaline medium provides higher number counter cations to release the silver ion from the treated fabric. The similar release behavior of silver ions has been observed by Yan *et al*. [23] that the concentration of silver ion release is higher in alkaline perspiration fluid compared to acidic and neutral perspiration solution.

Mechanical Properties of Finished Fabric

The mechanical properties of the silver intercalated clay finished fabric namely, tensile strength, elongation, bending length and crease recovery angle were evaluated and compared with the properties of untreated fabric. The results are summarized in Table 5. No significant change was found in tensile strength, bending length and elongation properties of the silver intercalated clay finished fabric as compared to untreated fabric. There is increase in the crease recovery angle (202 °) of treated fabric as compared to that of untreated fabric (187 °). This is due to the presence of DMDHEU as cross liking agent. This shows that the treatment does not affect the fabric mechanical properties besides imparting the durable antimicrobial finish along with crease resistant property. The bending length of finished fabrics decreases from 2 cm to 1.85 cm which show the fabric becomes slightly softer after finishing. The silver intercalated clay finished fabric showed light brown color. The color value of the silver intercalated clay finished fabric is shown in Table 6. The hue of the color increases with the concentration of the treatment. There are similar results are achieved by Hossam *et al.* that the color differences occurs while treatment of antibacterial agents to the fabrics [24].

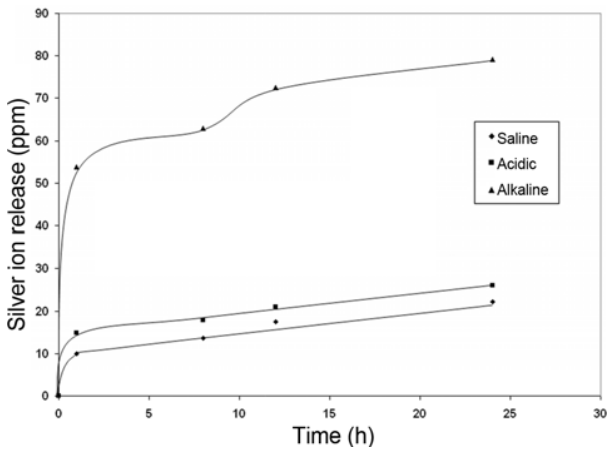


Figure 5. Silver ions release from fabrics with various perspiration medium using by ICP.

Table 5. Mechanical properties of fabrics in warp way

Properties	Unfinished fabric	Finished fabric
Weight add on (%)	-	0.8%
Tensile strength (breaking load) (N)	524 (7.69)	519 (5.26)
Extension (%)	7.36 (4.61)	8.22 (5.61)
Bending length (cm)	2.3 (4.15)	1.85 (7.82)
Crease recovery angle (degree)	187	202

Coefficient of variation is given in bracket.

Table 6. Color values of finished fabric

S. no	Concentration (%)	L*	a*	b*	c*	h°
Sample 1	0.25	85.35	0.97	3.52	3.65	74.55
Sample 2	0.50	79.81	2.78	7.38	7.88	69.38
Sample 3	1.00	62.89	8.71	25.48	26.93	71.13

Conclusion

The silver intercalated clay was prepared through ion exchange reaction with the silver loading of 10.5 wt %. The sliver intercalated clay is chemically attached with the cotton fabric along with DMDHEU crosslinking agent which was confirmed with FTIR studies. The silver modified fabric showed excellent antibacterial activity against *Staphylococcus aureus* and *E. coli* even after 10 washes which is equivalent to 50 home laundering. The silver ion releases from the fabric and inhibit the growth of the bacteria. The release of silver ion depends on the type of perspiration. Silver ion releases faster in alkaline perspiration as compared to acid

perspiration. This study shows that the silver intercalated clay has a great potential as an antimicrobial finishing agent for cotton textiles.

Acknowledgments

Authors are thankful to the M/S Lockheed & Martine, USA for the financial support for conducting the research work.

References

1. R. Purwar and M. Joshi, *AATCC Rev.,* **4**, 22 (2004).
2. M. Rai, A. Yadav, and A. Gade, *Biotechnol. Adv.*, **27**, 76 (2009).
3. Q. L. Feng, J. Wu, G. Q. Chen, F. Z. Cui, T. N. Kim, and J. O. Kim, *J. Biomed. Mater. Res.*, **52**, 662 (2000).
4. D. V. Parikh, J. V. Edwards, B. D. Condon, and A. D. Parikh, *AATCC Rev.*, **8**, 38 (2008).
5. N. Vigneshwaran, A. A. Kathe, P. V. Varadarajan, R. P. Nachane, and R. H. Balasubramanya, *J. Nanosci. Nanotechnol.*, **7**, 1893 (2007).
6. A. S. M. Raja, G. Thilagavathi, and T. Kannaian, *Indian J. Fibre Text. Res.*, **35**, 59 (2010).
7. A. I. Wasif and S. K. Laga, *Autex Res. J.*, **9**, 5 (2009).
8. S. Ghosh, S. Yadav, N. Vasanthan, and G. Sekosan, *J. Appl. Polym. Sci.*, **115**, 716 (2010).
9. B. Tomsic, B. Simoncic, B. Orel, M. Zerjav, H. Schroers, A. Simoncic, and Z. Samardzijia, *Carbohydr. Polym.*, **75**, 618 (2009).
10. I. Perlshtein, G. Applerot, G. G. Perkas, S. Mikhailov, and A. Gedanken, *Nanotechnology*, **19**, 245075 (2008).
11. B. Madheswaran, S. Shanmugam, and S. Tetsuo, *J. Coat. Technol. Res.*, **14**, 735 (2017).
12. S. Jiang, E. Newton, C.-W. M. Yuen, and C.-W. Kan, *Text. Res. J.*, **77**, 85 (2007).
13. E. Hontañón, J. Meyer, M. Blanes, V. Cambra, and X. Guo, *Inter. J. Theor. Appl. Nanotechnol.*, **4**, 17 (2016).
14. M. Alexandre and P. Dubois, *Mater. Sci. Eng.*, **28**, 1 (2000).
15. R. Purwar, K. Sai Goutham, and C. Mohan Srivastava, *Fiber. Polym.*, **17**, 1206 (2016).
16. S. M. Magaña, P. Quintana, D. H. Aguilar, J. A. Toledo, C. Ángeles-Chávez, M. A. Cortés, L. León, Y. Freile-Pelegrín, T. López, and R. M. Torres Sánchez, *J. Mol. Catal. A Chem.,* **281**, 192 (2008).
17. K. Malachová, P. Praus, Z. Pavlíčková, and M. Turicová, *Appl. Clay Sci.*, **43**, 364 (2009).
18. P. Praus, M. Turicova, and M. Valskova, *J. Braz. Chem. Soc.*, **19**, 549 (2008).
19. M. Joshi, R. Purwar, J. S. Udakhe, and R. Sreedevi, *U. S. Patent*, 9192625 (2015).
20. H. Nakashima, N. Miyano, and T. Takatuka, *J. Health Sci.*, **54**, 390 (2008).
21. A. S. Maryan, M. Montazerb, and T. Harifi, *Carbohydr. Polym.*, **115**, 568 (2015).
22. A. Hebeish, A. El-Shafei, and S. Sharaf, *Carbohydr. Polym.*, **103**, 442 (2014).
23. Y. Yan, Y. Haifeng, L. Junfang, L. Xiaojing, and W. Chao, *Text. Res. J.*, **82**, 1422 (2012).
24. E. E. Hossam and M. K. Zahran, *Int. J. Biol. Macromol.*, **75**, 106 (2015).

Arabian Journal for Science and Engineering
https://doi.org/10.1007/s13369-021-06017-4

RESEARCH ARTICLE-ELECTRICAL ENGINEERING

Application of Hybrid Metaheuristic Technique to Study Influence of Core Material and Core Trench on Performance of Surface Inset PMSM

Monika Verma[1] · **Mini Sreejeth**[1] · **Madhusudan Singh**[1]

Received: 12 April 2021 / Accepted: 14 July 2021
© King Fahd University of Petroleum & Minerals 2021

Abstract
Lightweight electric motor topologies used in electric transport applications require high-power density structures. Selection of suitable core material with appropriate physical geometry is important at design stage to reduce the core losses and the weight of the motor. This paper presents a novel hybrid metaheuristic optimization technique termed as Aquila-Grasshopper-Optimization (AGO), inspired by the hunting behavior of Aquila and swarming behavior of grasshoppers in search of the food source. This technique merges the merits of both Aquila optimizer (AO) and Grasshopper optimization algorithm (GOA) to create a balance between exploration and exploitation stages. By introducing the concept of normalization of the distances between the solutions (swarms) of the optimization problem at both narrowed and expanded exploration and exploitation stages of AO, the population is searched thoroughly to avoid the possibility of converging to any local optimum and to increase the convergence rate to reach the global optima. The performance of the proposed technique is first tested for various benchmark functions and thereafter implemented to the real-world problem to optimize the motor geometry of Surface Inset Permanent Magnet Synchronous Motor (SI-PMSM) to reduce the overall core losses. Two different core materials are employed in the core of the motor to study the influence of core material on the performance of the motor. In addition, the study is expanded by providing suitable trench cuts in motor core so that weight of the motor can be reduced substantially for lightweight motor applications. It is established that the variation in properties of core material can increase the power density and AGO technique can minimize core losses to obtain optimized motor core geometry thereby improving the overall electromagnetic performance of the motor.

Keywords Electrical steel · SI-PMSM · FEM · Power density · Core loss · Aquila optimizer · Grasshopper optimization

1 Introduction

Electric motors employed at the energy conversion stage of electric vehicle powertrain (EVP) systems have the most significant role in energy conservation of the EVP structure. Their performance factors can be easily scaled up or down based on the operating conditions. Research attempts for improving the performance of electrical machines using different fabrication techniques with appropriate core material have been reported. Research in the area of SI-PMSM for EV application is attracting a lot of interest [1]. The newly reported research areas include production technology of different height core rings with half-open slots for utilization in high-speed permanent magnet (PM) brushless DC motors [2], study of the performance of doubly salient PM machine with respect to two types of stator iron core segments with E shape and π shape design [3]; core loss calculation of Interior PM synchronous machine for space vector pulse width modulation excitation [4] etc.

In PM linear synchronous generator, core loss analysis using curve fitting method facilitates accurate study for slot less stator over wider frequency range [5]. This inspires to analyze the core loss modeling using curve fitting strategies effectively. In soft magnetic composite (SMC) electrical machines, core loss models have been developed using Poynting theorem by considering the loss coefficients which are dependent on frequency [6]. Core loss properties for stator and rotor cores composed of non-oriented (NO),

✉ Monika Verma
monikaverma_phd2k17@dtu.ac.in

[1] Department of Electrical Engineering, Delhi Technological University Rohini, Delhi, India

Published online: 13 August 2021

nano-crystalline magnetic materials (NMM) and NO-NMMs for a PMSM under inverter excitation [7] have also been reported. Core losses may be reduced using amorphous magnetic material stator core in place of NO steel core in PMSMs [8]. Better electromagnetic performance is achieved in partitioned stator PM machine models with dual E-core and C-core design [9]. The core loss modeling of PM motor is achievable by introducing loci of flux variation in different parts of the motor with the help of finite element transient analysis [10].

Finite element analysis can be a useful tool to study the core loss investigation in different parts of the motor. Finite section method can be used to model and separate core loss terms in electrical machines through extrapolation technique [11].

Core loss, being the considerable part of losses in motors, has to be reduced to increase efficiency of the motor, toward this due consideration has to be given at the design stage. The flux density and operating frequency have a major role in altering the values of core loss (hysteresis loss, eddy current loss and excess loss) [12]. But when it comes to motor geometry parameters, core loss shows big dependences over the variation in motor design parameters. This problem of minimizing the core loss through optimization of motor geometry under constraints of feasible range of design parameters has been analyzed and solved in this paper by introducing a hybrid AGO technique.

Considering specific drawbacks of the traditional optimization methods, the meta-heuristic (MH) techniques are gaining attention of researchers nowadays. These are nature inspired. There are majorly four categories of MH: Physics-based (PB), Swarm-based (SB), Evolution-based (EB) and Human-based (HB). SB methods bring in the behavior of swarms or animals [13–15]. EB uses processes inspired by theory of biological evolution containing the stages of reproduction, mutation or recombination and selection [16]. PB algorithms utilize the laws of Physics [17, 18]. Human behavior sets the inspiration for HB methods [19]. Apart from these, method like arithmetic optimization algorithm [20] inspires the researchers to follow the correct stepwise hierarchy to reach the target solution. It enhances the understanding of effects of no change policy in parameters of the algorithms during the process flow. MH techniques like hybrid antlion algorithms are not only useful in single objective problems but also show appreciable advantage in multi-objective task scheduling problems [21]. Hybridization in the basic methodology enhances the performance of group search optimizer (GSO) as compared to other basic MH techniques [22].

The two main phases of MH methods are (1) Exploration (also called diversification); which explores search zonal areas, and (2) Exploitation (also called intensification); which attempts to locate the optimum value from search zonal area. If the search area is not thoroughly analyzed from the entire population or the optimum point is not tested properly in the explored search area, there is a good chance that the overall solution may converge to local optima by missing out the global optimum. Hence, there is a need for maintaining the appropriate balance between the two phases so that the probability of skipping global optima is eliminated.

This paper addresses a real-world problem of optimizing motor geometry to obtain reduced core loss under specified constraints, by introducing a novel hybrid MH technique called Aquila-Grasshopper-Optimization (AGO). It blends the merits of Aquila optimizer (AO) and Grasshopper Optimization Algorithm (GOA).

Aquila is a wise and proficient hunter. AO method is designed on the basis of its hunting methods; (1) high ascend with a vertical sag called expanded exploration, (2) silhouette navigation with shortened slide called narrowed exploration, (3) low navigation with steady drop called expanded exploitation and (4) stretch and snatch prey called narrowed exploitation. It ensures that the optimization process invents the search area efficiently with the wide as well as narrow vision. However, it has few limitations such as imbalance between exploration and exploitation phases, possibility of convergence to local optima and disproportion in convergence pace in both phases [23].

Grasshoppers, on the other hand, come under the insect category that hop and fly like rolling cylinders forming a swarm. GOA is designed on the basis of social interactions of grasshopper swarms in search of the food source. This method is effective, robust, accurate and simple to implement. However, GOA has some disadvantages like slow speed to reach global optimum and wastage of iterations while converging to the new optimum (time consuming) [24]. In the proposed novel hybrid AGO technique, in both expanded and narrowed exploration phase, first the solutions are updated based on step (1) and step (2) of AO, respectively. Then in the updated current solution, the concept of normalizing the distance between solutions (based on GOA) in the specified range at each hunting step of AO is applied. Similarly, in both expanded and narrowed exploitation phase, first the solutions are updated on the basis of step (3) and Step (4) of AO, respectively, and then solutions are updated by normalizing the distance between the solution sets (based on GOA). Also, greedy selection is performed to update better fitness value of the solution at each step to obtain the global optima.

The major aspects covered in the reported research work include:

- A novel hybrid AGO mechanism is proposed by the authors by integrating the exploration and exploitation potentials of both AO and GOA.

- The proposed hybrid AGO algorithm is tested against twenty-three benchmark functions and its performance compared with the classical MH techniques.
- A real-world problem of minimizing the core loss by optimizing the motor geometry is implemented for SI-PMSM with different types of core materials.
- The weight of the motor is reduced by introducing trench cuts in the outer stator and inner rotor periphery of the motor. Its impact over the electromagnetic performance of the motor is investigated and reported.

In addition to the introduction, this paper is organized into four sections. Section 2 covers the formulation of core loss model in terms of motor geometry parameters. The proposed novel AGO algorithm is described in Sect. 3. Section 4 presents the results and discussions, followed by the conclusion in Sect. 5

2 Formulation of Core Loss Model in Terms of Motor Geometry Parameters

Following points illustrate the importance of the proposed research problem:

- For higher power density of PMSM, core loss is among the most critical components that influence motor's performance. The motor performance highly depends on the core material properties. This aspect makes this study a potential topic of research for enhancing the possibility of this motor's employment in EVP systems.
- Core materials have different electromagnetic properties. However, it evinces high dependence over the motor geometry parameters. Core loss modeling can be established in terms of motor geometrical parameters so that investigation can be carried out to minimize core losses in motor geometry.

A high kilo-watt (kW) low Revolutions Per Minute (RPM) model of 36 slots/14 poles SI-PMSM with distributed winding configuration is designed and analyzed. The motor design parameters are listed in Table 1. Trench cuts are provided in the motor geometry to obtain reduced weight and to study its effect on the electromagnetic performance of the motor. Designs of motor models are shown in Fig. 1, which illustrates: (1) stator and rotor without trench cuts, (2) stator and rotor with trench cuts and (3) stator with trench cuts and rotor without trench cuts. These models are analyzed for two different core materials to form six models listed in Table 2.

Models I and II are first analyzed with variations in (radial) magnet thickness (A), pole embrace (B) and stack length (C)

Table 1 Specifications of the Motor

Parameters	Value
Initial Motor Geometry	
Power	125 kW
Base Speed	3000 rpm
Frequency	350 Hz
Air gap length	0.5 mm
Stator outer diameter	200 mm
No of stator turns	300
Slot height	16.74 mm
Stator yoke thickness	22.76 mm
Rotor diameter	120 mm
Permanent Magnet	
Relative Permeability	1.09
Bulk Conductivity	625,000 Siemens/m
Remanence Flux Density	1.2 T
Magnetic Coercivity	890,000 A/m
Mass Density	7400 kg/m^3
Young's Modulus	147 Giga Pascals (GPa)

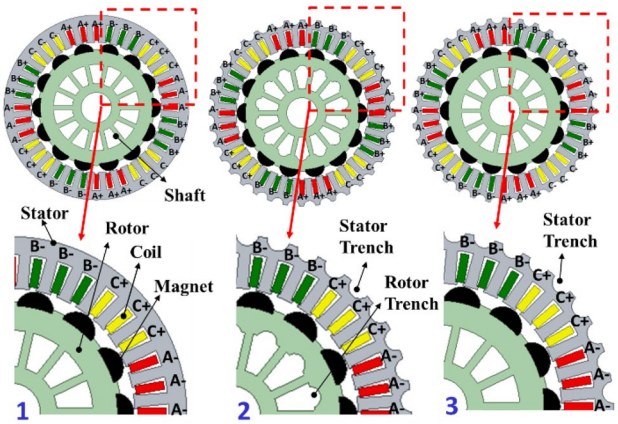


Fig. 1 Types of model designs

Table 2 Specification of Tested Models

Model no.	Core material used in the model	Presence of Trench Cuts in stator/rotor geometry Yes (Y) / No (N)
I	M19-29G	N
II	M400-50A1	N
III	M19-29G	Y
IV	M400-50A1	Y
V	M19-29G	Y (in stator), N (in rotor)
VI	M400-50A1	Y (in stator), N (in rotor)

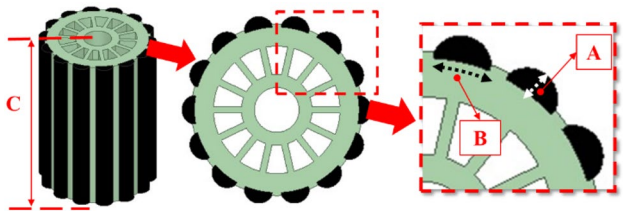


Fig. 2 Cross section of motor to present variable design parameters

as shown in Fig. 2. The core loss models, in terms of motor geometry parameters, are developed using finite element method (FEM) and surrogate modeling technique. In surrogate modeling or metamodeling, a relationship is established between the response of interest (that is core loss here) and specified input design parameters. For Design of Experiments (DOE), various combinations of design parameters are tabulated and corresponding FEM simulation responses are recorded. Thereafter, the coefficient matrix is computed using least square method (LSM) and coefficients of response surface models (RSM) [25, 26] are obtained to form the core loss models in terms of design variables of the motor.

The formulated Objective functions and constrains are given as follows:

Objective functions:

(1) Minimize $y_1 = -16.0281 - 11.0124A + 39.5326B + 1.7875C + 3.5295A.B - 0.0407A.C - 0.3057B.C + 0.642A^2 - 24.0917B^2 - 0.0009C^2$.

(2) Minimize $y_2 = 400.4877 - 37.93A + 247.695B - 4.004C - 13.3768A.B + 0.1171A.C - 3.3156B.C + 1.8694A^2 + 203.5722B^2 + 0.0208C^2$.

Constraints:

$3\text{mm} \leq A \leq 8\text{ mm}$,
$0.65 \leq B \leq 0.85$,
$150\text{mm} \leq C \leq 170\text{mm}$.

where y_1 and y_2 are the core losses of Models I and II.

The core material properties and corresponding loss curves are presented in Appendix A. The advantages of motor geometry selection, FEM simulations and the surrogate modeling analysis results are presented in Appendix B. The simulations are run for no load angle for avoiding the possible complexity of variation in core properties during loaded conditions.

3 Proposed AGO Algorithm

The proposed novel AGO Algorithm is presented in Fig. 3. The AGO algorithm begins with the stochastic generation of initial population of capable candidate solutions (X) between lower and upper bounds of the problem.

$$X = \begin{bmatrix} x_{1,1} & \cdots & x_{1,dim} \\ \vdots & \ddots & \vdots \\ x_{N,1} & \cdots & x_{N,dim} \end{bmatrix} \quad (1)$$

where current solutions X is generated using Eq. (2), $x_{N,dim}$ denotes the value of variable of Nth candidate solution located at (N, dim) in the matrix and dim denotes the dimension of the problem.

$$X_{ij} = lb_j + rand \times \left(ub_j - lb_j\right) \quad (2)$$

where $i = 1, 2, \ldots, N, j = 1, 2, \ldots\ dim$, $rand$ is random number, lb_j and ub_j are lower and upper bounds of jth variable.

This method takes into account the mathematical modeling of the hunting behavior of Aquila into four discrete steps. In each step, the traveling in search space is further carried out by including the swarming behavior of grasshoppers to reach the target food source with better fitness. The mathematical modeling of four hunting methods of Aquila is described as follows:

(1) High ascend with a vertical sag:

Initially, Aquila explores search space area from high ascend and chooses the best area for hunting the prey with a vertical sag. This behavior is modeled mathematically as follows:

$$X_1(t+1) = X_b(t) \times \left(1 - \frac{t}{T}\right) + \left(X_m(t) - X_b(t)\right) \times rand \quad (3)$$

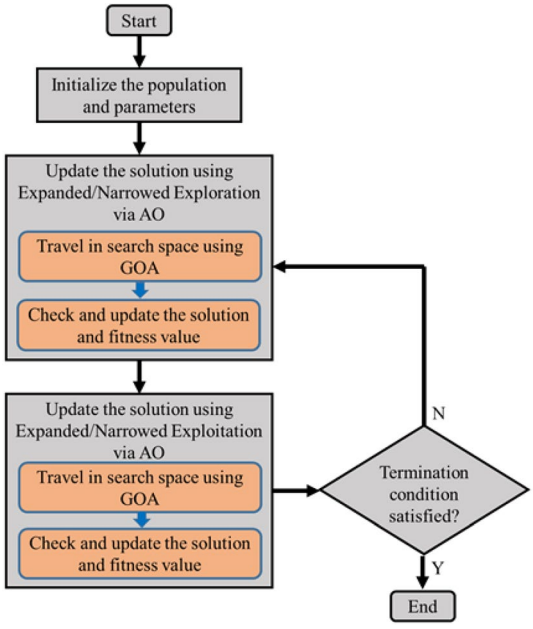


Fig. 3 Flow chart of AGO algorithm

where $X_1(t+1)$ is the updated solution at $(t+1)$th iteration, updated at step 1, $X_b(t)$ is the best updated solution up to current iteration, $\left(1-\frac{t}{T}\right)$ factor controls the exploration stage with respect to the change in iteration, T denotes maximum number of iterations, $X_m(t)$ is the mean of the solutions at tth iteration calculated using Eq. (4).

$$X_m(t) = \frac{1}{N}\sum_{i}^{N} x_{i,j}(t) \text{ for all } j = 1, 2, \ldots, dim \tag{4}$$

(2) Silhouette navigation with shortened slide:

Then, Aquila forms a circle above the prey food source, narrows the search area to attack. The method is termed as silhouette navigation with shortened slide. The mathematical model for this behavior is represented as:

$$X_2(t+1) = X_b(t) \times levy(d) + X_r(t) + (y - x) \times rand \tag{5}$$

where $X_2(t+1)$ is the updated solution at $(t+1)$th iteration updated at step 2, d is the dimension value of the problem, levy flight function is denoted by $levy(d)$ which is calculated using Eq. (6). $X_r(t)$ is the random solution selected from $[1, N]$ at tth iteration, y and x are representing contour spiral shape during the search as given in Eqs. (7), (8) and (9).

$$levy(d) = \frac{s \times u \times \sigma}{|v|^{1/\beta}} \tag{6}$$

$$\sigma = \frac{\Gamma(1+\beta) \times \sin\frac{\pi\beta}{2}}{\Gamma\left(\frac{1+\beta}{2}\right) \times \beta \times 2^{\frac{\beta-1}{2}}} \tag{7}$$

$$y = r \times \cos\theta \tag{8}$$

$$x = r \times \sin\theta \tag{9}$$

$$r = r_1 + p \times d_1 \tag{10}$$

$$\theta = -\omega \times d_1 + \theta_1 \tag{11}$$

where s, β, r_1, p, ω and θ_1 are constants. u and v are random numbers in [0,1]. d_1 is a vector of integers containing elements from 1 to dim.

(3) Low navigation with steady drop:

As the area of the prey is stated, Aquila navigates a bit low with steady drop. Here, prey area is being exploited with a wide vision. The mathematical model is represented by Eq. (12).

$$X_3(t+1) = (X_b(t) - X_m(t)) \times \alpha - rand + ((ub - lb) \times rand + lb) \times \delta \tag{12}$$

where $X_3(t+1)$ is the updated solution at $(t+1)$th iteration updated at step 3, α and δ are exploitation control parameters.

(4) Stretch and snatch prey:

When prey is close to Aquila, it stretches and snatches the prey in accordance with the random motion of the prey. Equation (13) represents the mathematical model for this behavior.

$$\begin{aligned} X_4(t+1) = QF \times X_b(t) - \left(g_1 \times X(t) \times rand\right) \\ - \left(g_2 \times levy(d)\right) + rand \times g_1 \end{aligned} \tag{13}$$

where $X_4(t+1)$ is the updated solution at $(t+1)$th iteration updated at step 4, QF is quality function calculated at tth iteration, g_1 and g_2 are the control parameter to track the prey by Aquila as given in (14), (15) and (1616).

$$QF(t) = t^{\frac{2\times rand()-1}{(1-T)^2}} \tag{14}$$

$$g_1 = 2 \times rand - 1 \tag{15}$$

$$g_2 = 2 \times \left(1 - \frac{t}{T}\right) \tag{16}$$

The updated solutions at each step mentioned above are treated as new grasshopper positions at tth iteration. The positions of grasshoppers are then updated using the mathematical model of their social swarming behavior, represented by Eq. (17).

$$X_i = c\left(\sum_{\substack{j=1 \\ j \neq i}}^{N} c\frac{ub - lb}{2} s(\left|x_j - x_i\right|)\frac{x_j - x_i}{d_{ij}}\right) + \widehat{T}_d \tag{17}$$

where X_i denotes the position of ith grasshopper in d dimensional search space, given as: $X_i = \left(x_{i1}, x_{i2}, \ldots, x_{id}\right)$, c is a control coefficient calculated using Eq. (18), s calculates the social force strength using Eq. (19), d_{ij} is the distance between ith grasshopper x_i and jth grasshopper x_j as given in Eq. (20).

$$c = c\max - t \times \frac{c\max - c\min}{T} \tag{18}$$

$$s\left(\left|x_j - x_i\right|\right) = fe^{\frac{-\left|x_j - x_i\right|}{l}} - e^{-\left|x_j - x_i\right|} \tag{19}$$

$$d_{ij} = \left|x_j - x_i\right| \tag{20}$$

where f and l are attraction intensity factor and length scale factor of attraction, respectively. $c\max$ and $c\min$

are maximum and minimum value of c coefficient which controls the search area based on the iterations [27]. The pseudo-code of the proposed algorithm is shown in Fig. 4.

4 Results and Discussion

This section enumerates the performance analysis of the proposed novel AGO algorithm under two sub-heads (a) Testing of AGO algorithm on benchmark Functions and (b) Validation of AGO algorithm on a real-world problem of optimizing motor geometry to obtain reduced core loss. The performance of proposed algorithm is compared with classical MH techniques.

4.1 Testing of AGO Algorithm on Benchmark Functions

The proposed algorithm is implemented in MATLAB 2016b and tested for various benchmark functions. The definitions

```
Initialize the population X
Initialize the parameters of AO and GOA
While (end criteria not satisfied)
   store best solution X_b(t)
   calculate fitness value
   for each solution (i=1,2,...N)
      update mean value of current solution X_m(t)
      update parameters c, x, y, g_1, g_2, levy(d)
      if t ≤ (2/3) × T
         %% Step (1) Expanded exploration
            update solution using Eq. (3)
            check and update fitness
            normalize the distance between solutions
            update position using Eq. (17)
            check and update fitness
         %% Step (2) Narrowed exploration
            update solution using Eq. (5)
            check and update fitness
            normalize the distance between solutions
            update position using Eq. (17)
            check and update fitness
      else
         %% Step (3) Expanded exploitation
            update solution using Eq. (12)
            check and update fitness
            normalize the distance between solutions
            update position using Eq. (17)
            check and update fitness
         %% Step (4) Narrowed exploitation
            update solution using Eq. (13)
            check and update fitness
            normalize the distance between solutions
            update position using Eq. (17)
            check and update fitness
      end if
   end for
   t = t + 1
end while
Return the best solution and best fitness
```

Fig. 4 The pseudo-code of proposed AGO algorithm

of benchmark functions and the comparative study with other MH techniques are shown in Table 3. Various unimodal, multimodal and fixed dimension multimodal benchmarks functions are used for the comparative analysis of the performance of proposed AGO algorithm with classical MH techniques like Artificial Bee Colony (ABC) method, Particle Swarm Optimization (PSO) method, Teaching Learning-Based Optimization (TLBO) method and Differential Evolution (DE) method.

The simulations are run for 1000 iterations with 20 solutions (search agents). The values of the computational parameters used in different algorithms are listed in Table VII. The algorithms are evaluated and are awarded ranks based on their successful convergence to the optima for different functions. The exploitation tendency of AGO is tested for unimodal functions F1-F7, and the exploration tendency of AGO is tested for multimodal functions F8-F13. On the basis of the number of functions for which a particular algorithm shows the successful least amount of fitness to the function, ranks are awarded to them.

It is observed that AGO produces the fittest value with the ongoing process flow of the iterations securing rank 1. The performance of AGO is found to be better than AO and GOA, which secure ranks 2 and 3, respectively. Whereas, PSO and TLBO are ranked 4 and 5 respectively. DE did not convergence for functions F7, F8, F12, F13 and secured rank 6. While ABC operates satisfactorily for few functions like F1-F5, F9, F10, F11 it failed to converge to the required minima for functions like F6, F7, F8, F12, F13 thus, securing rank 7.

The performance of AGO is tested for the low dimension problems via functions F14-F23. AGO is found to be converging to the optimum value in fewer iterations in almost all the functions F14-F23 and hence secures rank 1. Similarly, other algorithms are also awarded the ranks based on their corresponding performance. The balance between exploration and exploitation carried out by AGO is shown in Fig. 5 for different benchmark functions. It is evident from the convergence curve of AGO (Fig. 5) that the qualitative results are approaching the target by searching the population thoroughly for unimodal functions whereas the results are following stepwise exploration or exploitation based on the complexity of the multimodal functions. AGO quickly experiences the transition from exploration to exploitation in case of fixed dimensional multimodal functions. The convergence curves for different algorithms are compared in Fig. 6 to test the acceleration tendency of AGO. It is observed that for unimodal functions F1-F7, AGO experiences a favorable drop off rate as compared to other algorithms. For multimodal functions F8-F23, AGO shows appreciable deceleration rate and hence shows the rapid convergence capability. While the other algorithms show the drawback

Table 3 Definitions of benchmark functions and comparative analysis results

Function		Description	Range	f_{min}	AGO	AO	GOA	ABC	PSO	TLBO	DE
Unimodal	F1	$f(x) = \sum_{i=1}^{n} x_i^2$	[−100,100]	0	8.839E−35	1.770E−19	2.509E−06	4.507E−11	1.715E−25	3.854E−33	8.383E−17
	F2	$f(x) = \sum_{i=1}^{n} \lvert x_i \rvert + \prod_{i=0}^{n} \lvert x_i \rvert$	[−10,10]	0	9.111E−23	1.554E−10	0.283E+00	2.879E−16	1.532E−12	1.020E−17	6.647E−06
	F3	$f(x) = \sum_{i=1}^{d} \left(\sum_{j=1}^{i} x_j \right)^2$	[−100,100]	0	3.831E−53	1.044E−19	1.633E−04	2.448E−10	2.912E−24	5.817E−32	1.234E−17
	F4	$f(x) = \max_i \{ \lvert x_i \rvert, 1 \le i \le n \}$	[−100,100]	0	2.521E−62	9.034E−10	1.968E−04	5.199E−17	4.850E−12	4.703E−16	2.874E−07
	F5	$f(x) = \sum_{i=1}^{n-1} \left[100 \left(x_i^2 - x_{i+1} \right)^2 + \left(1 - x_i \right)^2 \right]$	[−30,30]	0	8.222E−04	8.306E−04	1.011E+00	0.401E+00	0.000E+00	0.000E+00	0.000E+00
	F6	$f(x) = \sum_{i=1}^{n} \left(\lvert x_i + 0.5 \rvert \right)^2$	[−100,100]	0	2.152E−25	4.556E−06	1.012E−09	0.500E+00	0.000E+00	4.264E−08	1.080E−−08
	F7	$f(x) = \sum_{i=0}^{n} i x_i^4 + random[0,1)$	[−128,128]	0	5.318E−03	2.661E−03	5.812E−03	0.098E+00	3.292E−05	0.004E+00	0.011E+00
Multi-modal	F8	$f(x) = \sum_{i=1}^{n} \left(-x_i \sin \left(\sqrt{\lvert x_i \rvert} \right) \right)$	[−500,500]	−418.982	−413.586	−412.784	−417.336	−402.968	−408.229	−400.868	−362.886
	F9	$f(x) = \sum_{i=1}^{n} \left[x_i^2 - 10 \cos \left(2\pi x_i \right) + 10 \right]$	[−5.12,5.12]	0	1.294E−10	0.000E+00	4.682E+00	4.052E−10	0.000E+00	0.000E+00	3.979E+00
	F10	$f(x) = -20 \exp \left(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^{n} x_i^2} \right) - \exp \left(\frac{1}{n} \sum_{i=1}^{n} \cos \left(2\pi x_i \right) \right) + 20 + e$	[−32,32]	0	1.764E−16	8.881E−16	4.619E−05	4.148E−16	8.881E−16	8.881E−16	6.525E−06
	F11	$f(x) = 1 + \frac{1}{4000} \sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos \left(\frac{x_i}{\sqrt{i}} \right)$	[−600,600]	0	1.775E−15	0.000E+00	0.096E+00	2.390E−05	0.000E+00	4.440E−15	0.012E+00
	F12	$f(x) = \frac{\pi}{n} \left\{ 10 \sin \left(\pi y_1 \right) \right\} + \sum_{i=1}^{n-1} \left(y_i - 1 \right)^2 \left[1 + 10 \sin^2 \left(\pi y_{i+1} \right) + \sum_{i=1}^{n} u \left(x_i, 10, 100, 4 \right) \right]$; where $y_i = 1 + \frac{x_i + 1}{4}, u \left(x_i, a, k, m \right) = \begin{cases} K \left(x_i - a \right)^m; & \text{for } x_i > a \\ 0; & \text{for } -a \le x_i \le a \\ K \left(-x_i - a \right)^m; & \text{for } x_i \le -a \end{cases}$	[−50,50]	0	1.498E−06	1.505E−05	5.519E−06	1.000E+00	4.711E−03	4.326E−04	3.110E+00
	F13	$f(x) = 0.1(\sin^2 \left(3\pi x_1 \right) + \sum_{i=1}^{n} \left(x_i - 1 \right)^2 \left[1 + \sin^2 \left(3\pi x_i + 1 \right) \right] + \left(x_n - 1 \right)^2 \left[1 + \sin^2 \left(2\pi x_n \right) \right) + \sum_{i=1}^{n} u \left(x_i, 5, 100, 4 \right)$	[−50,50]	0	1.537E−07	1.353E−06	1.327E−06	1.000E+00	1.349E−02	6.816E−03	0.019E+00
			Rank		1	2	3	7	4	5	6

Table 3 (continued)

Function		Description	Range	$f_{\min}$	AGO	AO	GOA	ABC	PSO	TLBO	DE
Fixed Dimensional Multimodal	F14	$f(x) = \left(\frac{1}{500} + \sum_{j=1}^{25}\left(j + \left(x_1 - a_{1j}\right)^6 + \left(x_2 - a_{2j}\right)^6\right)^{-1}\right)^{-1}$	[−65,65]	1	0.999	0.996	1.992	0.978	0.998	0.998	12.670
	F15	$f(x) = \sum_{i=1}^{11}\left[a_i - \frac{x_1\left(b_i^2 + b_i x_2\right)}{b_i^2 + b_i x_3 + x_4}\right]^2$	[−5,5]	0.00030	7.730E−04	3.320E−04	7.034E−04	3.100E−02	2.989E−03	2.972E−02	3.321E−01
	F16	$f(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4$	[−5,5]	−1.0316	−1.031	−1.031	−1.031	−1.019	−1.002	−1.029	−1.006
	F17	$f(x) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos x_1 + 10$	[−5,5]	0.398	0.398	0.397	0.397	0.397	0.397	0.377	1.943
	F18	$f(x) = \left[1 + \left(x_1 + x_2 + 1\right)^2\left(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2\right)\right] \times \left[30 + \left(2x_1 - 3x_2\right)^2\left(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2\right)\right]$	[−2,2]	3	3	3.001	3	2.998	2.909	2.998	2.990
	F19	$f(x) = -\sum_{i=1}^{4} c_i \exp\left(-\sum_{i=1}^{3} a_{ij}\left(x_j - p_{ij}\right)^2\right)$	[−1,2]	−3.86	−3.861	−3.860	−3.857	−1.899	−1.899	−1.899	−1.899
	F20	$f(x) = -\sum_{i=1}^{4} c_i \exp\left(-\sum_{i=1}^{6} a_{ij}\left(x_j - p_{ij}\right)^2\right)$	[0,1]	−3.32	−3.322	−3.304	−3.322	−1.169	−1.169	−1.166	−1.169
	F21	$f(x) = -\sum_{i=1}^{5}\left[\left(X - a_i\right)\left(X - a_i\right)^T + c_i\right]^{-1}$	[0,1]	−10.1532	−10.055	10.151	−5.055	−8.152	−10.001	−10.113	−5.101
	F22	$f(x) = -\sum_{i=1}^{7}\left[\left(X - a_i\right)\left(X - a_i\right)^T + c_i\right]^{-1}4$	[0,1]	−10.4028	−10.402	−10.400	−10.402	−9.908	−10.376	−10.400	−5.128
	F23	$f(x) = -\sum_{i=1}^{10}\left[\left(X - a_i\right)\left(X - a_i\right)^T + c_i\right]^{-1}4$	[0,1]	−10.5363	−10.427	−10.535	−10.536	−9.693	−10.359	−10.519	−6.529
			Rank		1	2	3	6	4	5	7

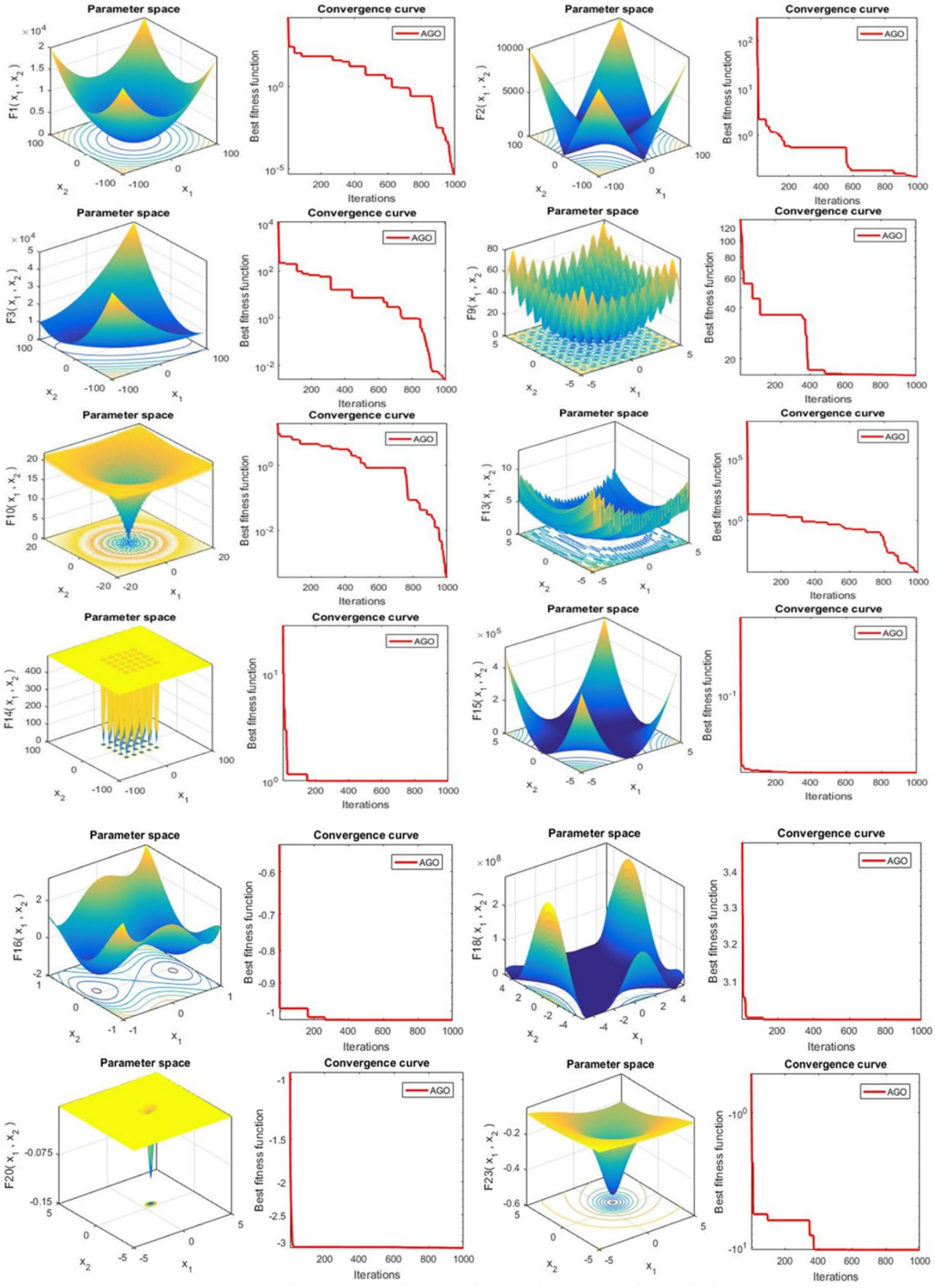


Fig. 5 Performance of AGO for various benchmark functions

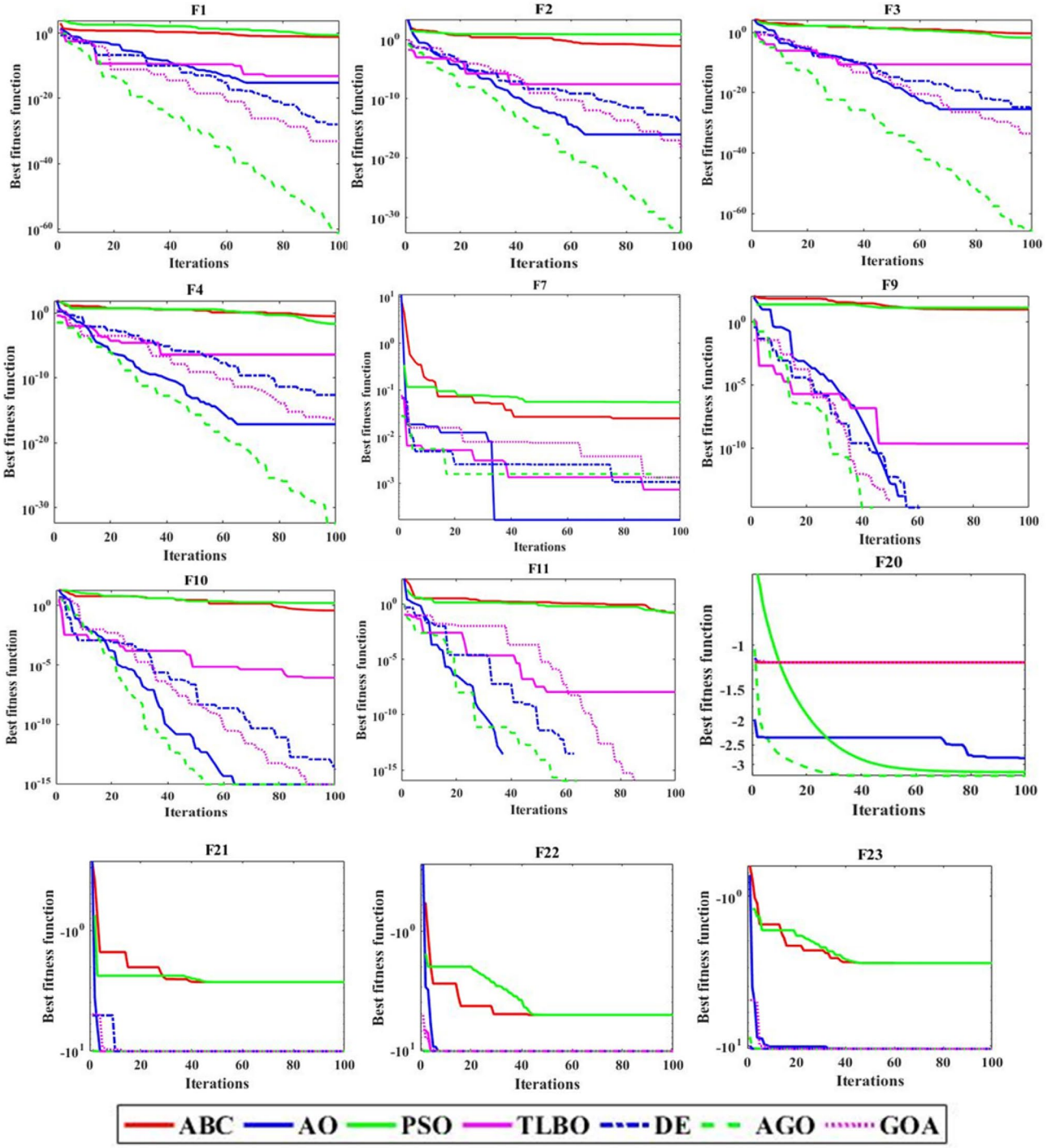


Fig. 6 Comparative analysis results for convergence test of different algorithms

Table 4 Performance analysis results of algorithms for core loss models y_1 and y_2

Method	y_1		Rank	y_2		Rank
	(A,B,C)	y_{min}		(A,B,C)	y_{min}	
AGO	(8,0.85,150)	136.216	1	(8,0.85,150)	68.529	1
AO	(7.9,0.85,150)	137.607	3	(7.2,0.85,150)	71.251	4
GOA	(7.77,0.84,151.34)	139.523	7	(8,0.82,154.56)	71.337	5
ABC	(7.97,0.85,150)	137.297	2	(8,0.85,156.05)	71.437	6
PSO	(8,0.82,150)	137.637	4	(7.82,0.81,150)	69.583	2
TLBO	(7.93,0.75,150)	139.051	6	(7.88,0.79,150.29)	69.958	3
DE	(7.6,0.85,150)	138.836	5	(6.4,0.79,150)	76.135	7

of converging to local optima like in F1, F2, F3, F4, F9, F10, F11 and undergo sluggishness in reaching the optima such as in F7, F20, F21, F22, F23.

4.2 4.2 Validation of AGO Algorithm on a Real-World Problem of Optimizing Motor Geometry to Obtain Reduced Core Loss

The developed AGO algorithm is applied on the core loss model formulated in Sect. 2. The results of the analysis for the loss models are presented in Table 4. The algorithms are ranked based on the fitness values of y_1 and y_2. The proposed AGO algorithm gives the least values of core loss models of SI-PMSM as compared with the other algorithms.

At rated speed of 3000 rpm, it is observed that Models I and II exhibit the least core losses of 136.2 mW and 68.5 mW, respectively (for A = 8 mm, B = 0.85 and C = 150 mm). For better motor performance, output torque should be high and torque ripple should be lower as compared to the initial models. The performance of two motors in remaining combinations of variables is not considered for further analysis since values of torque and core losses from the motor contradict. Further to optimize effective magnetic path length, stator yoke length, t_{sy} is varied radially and corresponding torque and torque ripple plots are presented in Fig. 7. It is observed that torque increases in direct proportion to t_{sy} and torque ripple decreases. The torque and ripple torque are retained stable after t_{sy} = 10.8 mm. To meet the torque requirement and to satisfy the objective of using less core material, t_{sy} is fixed to 10.3 mm, thereby obtaining the torque = 412.7 Nm and torque ripple = 0.13 pu.

The optimized geometry for all six models (mentioned in Table 2) is analyzed through simulation studies and torque, torque ripple, core loss in stator and core loss in rotor of

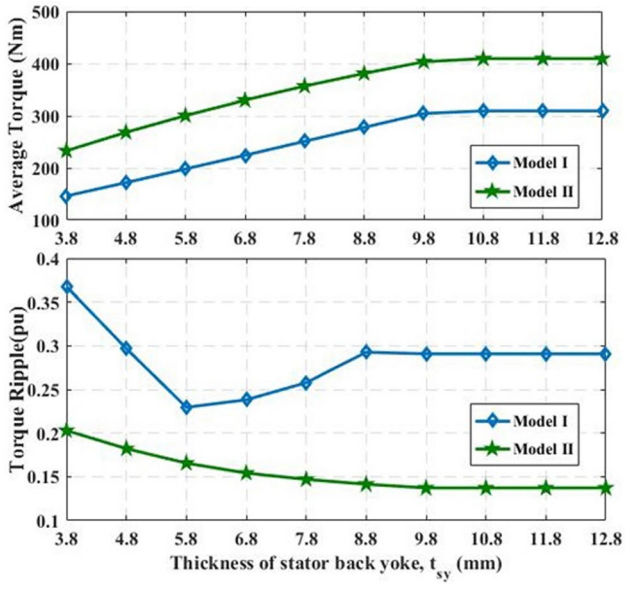


Fig. 7 Plot of torque and torque ripple vs variable stator yoke length (for torque ripple, per unit is abbreviated as pu)

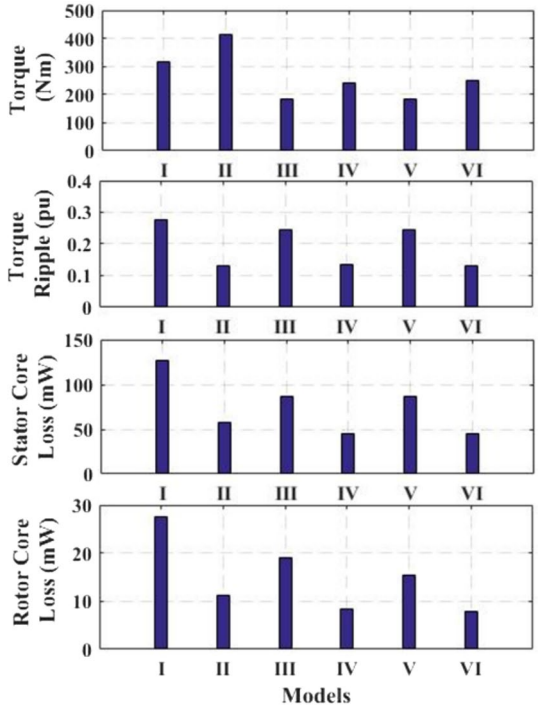
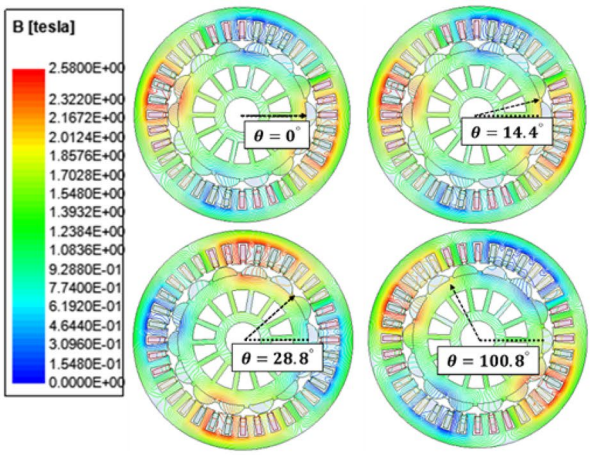


Fig. 8 Performance comparison of six models

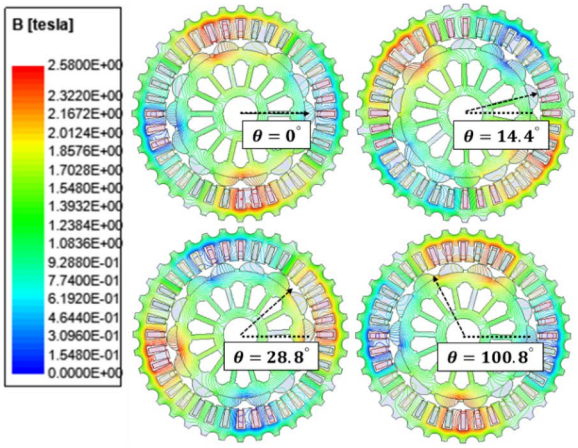
the different models are compared. Figure 8 presents the corresponding comparison of six models when operated at rated speed of 3000 rpm and 17A of current. It is observed that Models III and V produce equal average torque of 182 Nm and torque ripple of 0.24 pu. Also, Models IV and VI produce equal average torque of 248 Nm and torque ripple of 0.21 pu. Thus, it is inferred that the presence of trenches in rotor core does not have any significant effect on motor torque performance but this additional feature is helpful in reducing motor weight which is presented in Fig. 10. Models I and II produce average torque of 317 Nm and 412 Nm, respectively. Core losses in stator produced by Models I and II are 126.4 mW and 57.2 mW, while Core losses in rotor are 27.4 mW and 11.3 mW, respectively. Torque ripple produced by Model II is 0.13 pu which is 53% lesser than that produced by Model I. Model II produces the highest average torque of 412.7 Nm and the least torque ripple of 0.13 pu.

The magnetic flux density distributions in Models II and IV at different time instants are shown in Fig. 9 (a) and (b), respectively. On outer periphery of the stator yoke and inner periphery of the rotor yoke, magnetic flux density is approximately 0 T and 0.4 T, respectively, at different time instants depending upon the rotor position θ. In Fig. 9 (c), the partial motor models are presented to understand the influence of presence of trench cuts in stator and rotor yokes with respect to magnetic flux density distribution. Points P_{s1}, P_{s2}, P_{s3}, P_{s4}, P_{s5} in stator and points P_{r1}, P_{r2} in rotor are representing the points of observation (trench cuts positions) in partial models, respectively. It is observed that if in Model II, trench cuts are provided in stator and rotor yoke geometry at these points of observation, magnetic flux density gets distributed in the confined yoke geometry near to the slots. Thus, Model II exhibits least core losses of 68.5 mW. In Model IV, the core material is reduced without affecting the flux density and hence lightweight motor structure is achieved, which is suitable for electrical transportation and aerospace applications. The weight and power density of the six models are compared in Fig. 10.

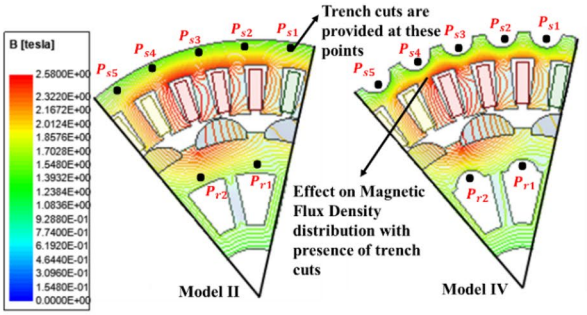
For experimental validation of models, the torque performances of optimized motor prototype of model I and Model II are tested and results are compared in Fig. 11. It can be observed that Model II exhibits reduction in torque ripple by 53% as compared to Model I at rated parameters. Output torque from Model II is higher than that of Model I. The torque speed characteristics along with the efficiency plots are presented in Fig. 12. It is observed that both Models I and II perform satisfactorily. Model II shows enhancement in efficiency of approximately 2.5%. Therefore, using non-grain-oriented core material in stator and rotor core improves the overall performance of the motor. The simulation and experimental analysis confirm the validity of this work.



(a) Magnetic Flux Density for model II



(b) Magnetic Flux Density for model IV



(c) Partial motor diagrams at *t=0 msec (θ=0°)*

Fig. 9 Magnetic Flux Densities at $t=0$ ms ($\theta=0°$), $t=0.8$ ms ($\theta=14.4°$), $t=1.6$ ms ($\theta=28.8°$) and $t=5.6$ ms ($\theta=100.8°$)

Furthermore, sensitivity analysis of core loss and efficiency is carried out for Models I and II by introducing deviations in design parameters $\boldsymbol{p} = [\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{t_{sy}}]$.

The deviations are applied to one parameter at a time while keeping other variables unchanged, as given as follows:

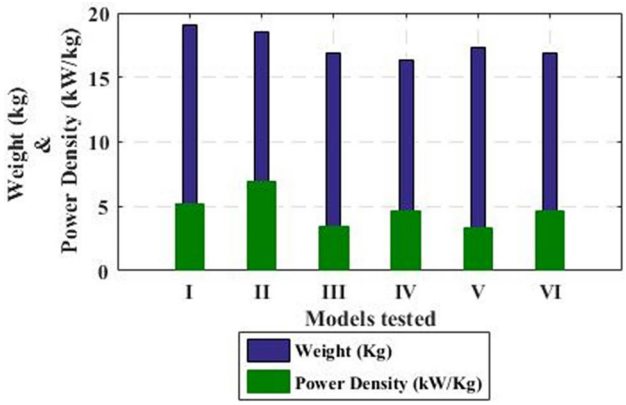


Fig. 10 Comparison of Weight and Power Density of six models

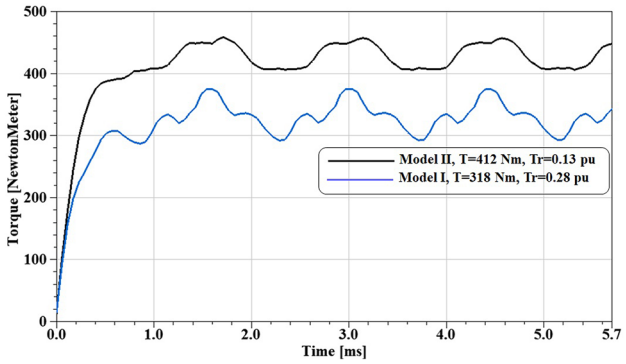


Fig. 11 Torque ripple comparison of Model I and II

$$S_j^d = \frac{d(p_1, \ldots, p_j + \Delta p_j, \ldots, p_n) - d(p_j)}{\Delta p_j} \tag{21}$$

where d denotes the data function; core loss or efficiency and p_j denotes the jth independent variable. Figure 13 shows the core loss deviation (%) as a function of parameter deviation (%). Table 5 presents the sensitivity values for each parameter.

It is observed that core loss and efficiency are highly sensitive to magnet thickness (A), while deviation in pole embrace (B), Stack length (L) and stator yoke length (t_{sy}) has less effect on model performance. Also, Model II is comparatively less sensitive to parameters disturbance than Model I. The geometrical and electrical disturbances also develop undesired behavior in the motor, like fault occurrence due to unknown disturbances and unaligned eccentricity of rotor. Such sensitivity analysis proposes an interesting approach to further analyze the motor performance and develop a robust and reliable motor designs [28, 29]. This can be considered as a future task to enhance knowledge of the subject with different core materials.

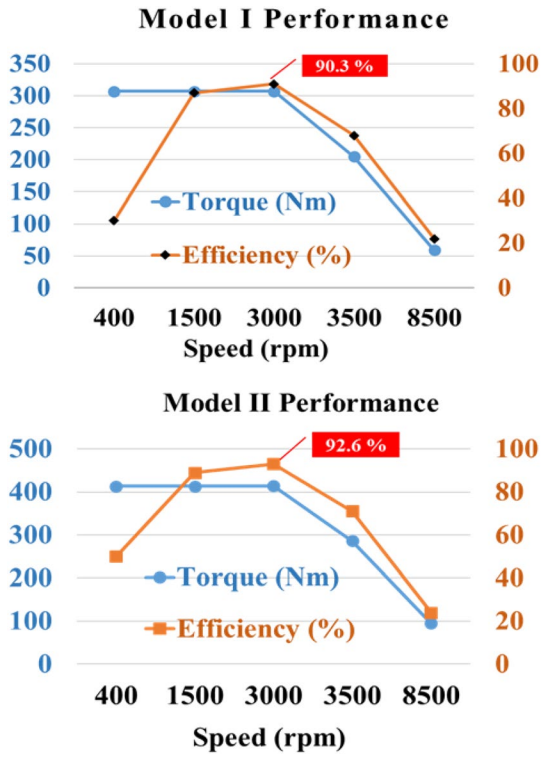


Fig. 12 Torque speed characteristics of Model I and II

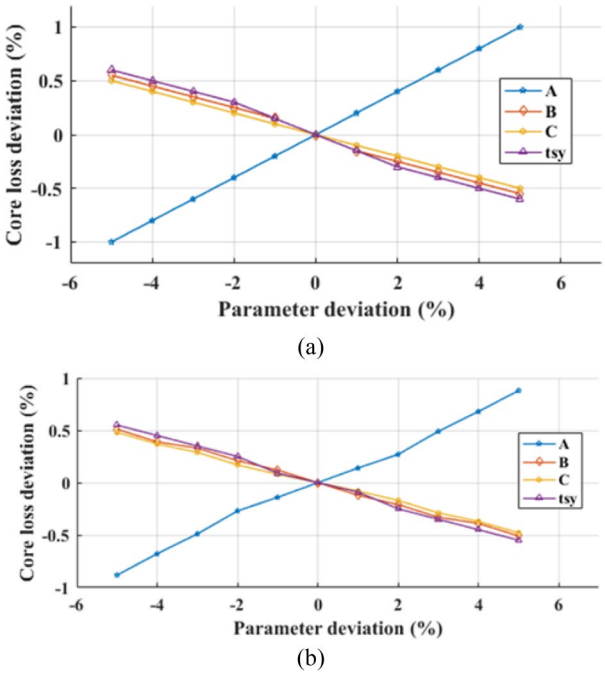


Fig. 13 Sensitivity analysis (**a**) Model I (**b**) Model II

Table 5 Data function sensitivity computation

	A	B	C	t_{sy}
Core loss				
Model I	0.80	0.45	0.40	0.50
Model II	0.68	0.39	0.37	0.45
Efficiency				
Model I	0.091	0.034	0.049	0.066
Model II	0.099	0.020	0.037	0.048

5 Conclusion

The research achievements reported in this paper are summarized as follows:

- This paper presents a novel hybrid metaheuristic optimization technique termed as Aquila-Grasshopper-Optimization. The application of this hybrid metaheuristic optimization technique is elucidated to study the influence of core material and core trench on the performance of SI-PMSM.
- The core loss models in terms of motor geometry parameters are developed using FEM and surrogate modeling technique using different core materials. Six structurally different FEM models of 36 slot/14 pole SI-PMSM are designed and optimized to obtain high torque with less core losses.
- The performance of proposed novel AGO technique is tested for various benchmark functions and further implemented for the real-world problem of design optimization of SI-PMSM to achieve minimum core losses, through the use of optimized design parameters.
- It is shown that by optimizing motor geometry at design stage, power density of the motor can be increased. It is observed that usage of non-grain-oriented fully processed electrical steel in stator and rotor core of electric motor is one of the measures to improve the efficiency of SI-PMSM motor. Use of M400-50A1 in stator and rotor core results in enhancement in efficiency of approximately 2.5% and reduction in torque ripple by 53% as compared to M19-29G.
- It is established that appropriately designed trench cuts can be employed in stator and rotor structure without affecting the magnetic flux density for electric vehicle and aerospace applications, where lightweight, high torque motors are preferred.
- In addition, the model performance is investigated under deviation in parameter values using sensitivity analysis. The results show that Model with non-grain-oriented electrical steel is less sensitive to perturbations in parameters value.

Appendix A: Core Material Properties and Parameters of Conventional Optimization Algorithms

The electrical steel availed in PM machines is an iron alloy forged to exhibit certain magnetic properties. It is designed with small hysteresis area and high permeability, which results in low power loss per cycle. Magnetization curves of a low carbon silicon sheet (M19-29G) and a non-grain-oriented electrical steel (M400-50A1) are shown in Fig. 14. Iron loss curves of the two materials at different frequencies are presented in Fig. 15. As per European Standard EN 10,106, core materials are magnetic age free.

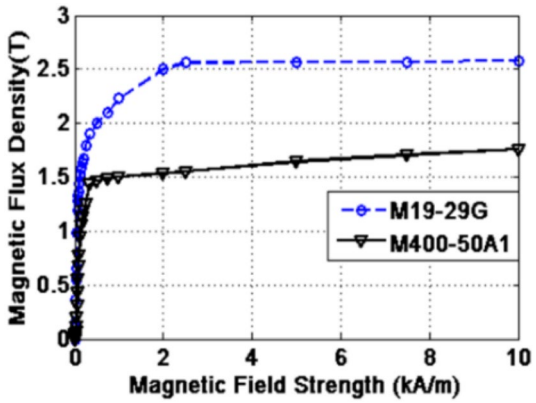


Fig. 14 B-H Curves of core materials

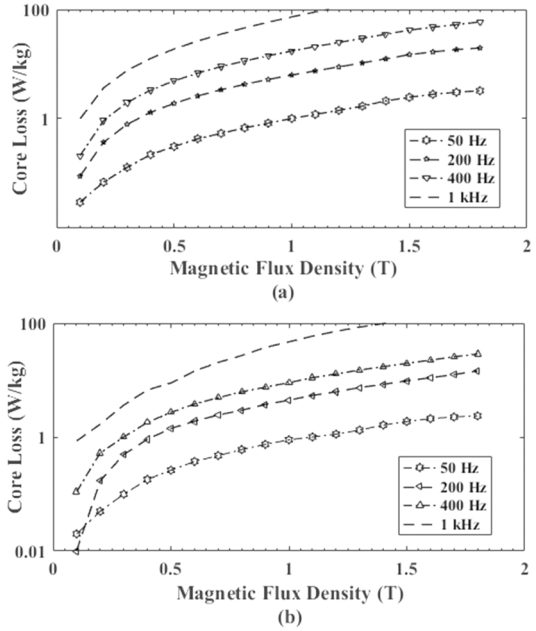


Fig. 15 Core Loss Curves of (**a**) M19-29G and (**b**) M400-50A1 at four different frequencies

By applying different range of compressive and tensile stresses, B-H loop of non-oriented electrical steel gets deformed at lower frequency ranges up to 100 Hz [30]. Also, the extrapolation of hysteresis loss from frequency measurement at 2 to 5 Hz results in erroneous segregation of excess power loss. In [31], T. J. Miller illustrates that the stresses like electrical stress and mechanical stresses are caused by direct on line motors. The detrimental effects of such stresses can be waived off by designing adjustable speed drives with controlled acceleration. In this study, the motor is analyzed at the rated operating conditions of 3000 rpm and 350 Hz. It is observed in [32] that hysteresis loss, eddy current loss and excess losses can be reduced by shaping appropriate domain structure of the core material. Here, external stress free condition is considered for the analysis in this work. Table 6 shows the specific properties of the two core materials (Table 7).

Table 6 Core Material Properties

Material	M19−29G	M400-50A1
Mass Density	7872 kg/m^3	7650 kg/m^3
Thickness	0.35 mm	0.50 mm
Composition	Fe–C (C < ~0.1%)	Fe-Si (2–3.5%Si)
Anisotropy of Loss	±17%	±10%
Resistivity	0.52 μΩ m	0.42 μΩ m
Hysteresis Loss Coefficient, k_h	2.05e−2 W/m^3	9.59e−3 W/m^3
Eddy Current Loss Coefficient, k_e	3.46e−5 W/m^3	3.78e−5 W/m^3
Excess Loss Coefficient, k_{exc}	5.96e−4 W/m^3	2.01e−5 W/m^3
B (5000 A/m)	2.57 T	1.64 T
Saturation Flux Density	2.58 T	1.7 T

Table 7 Values of computational parameters of different algorithms

Method	Parameter	Value
ABC	permissible no. of failures	5
PSO	Inertia weight	0.5
	Cognitive and social coefficients	1.5, 1.5
TLBO	Population size	20
DE	Crossover probability	0.8
	Scaling factor	0.85
GOA	cmax	1
	cmin	0.00004
	Attraction intensity factor	0.5
	Length scale factor of attraction	1.5
AO	Exploitation control parameters	0.1, 0.1
	ω	0.005
	s	0.01
	u	0.0265
	β	1.5
	θ_1	$3\pi/2$

The total core loss P in PMSM machine is the combination of three loss factors, i.e., hysteresis loss P_h, eddy current loss P_e and excess loss P_{exc}, given by:

$$P = P_h + P_e + P_{exc}$$

$$P = k_h f B^n + 2k_e (fB)^2 + k_{exc} (fB)^{3/2} \tag{22}$$

where f is the frequency, B is the peak flux density, k_h, k_e and k_{exc} are loss coefficients for hysteresis, eddy current and excess losses respectively, n is Steinmetz Constant, may vary from 1.5 to 2.5 depending on the material.

Appendix B: Advantages of Motor Geometry Selection, FEM Simulations and Surrogate Modeling Analysis Results

FEM modeling of complex structures allows the user to determine how critical factors like geometrical parameters and material properties, etc., affect the model performance with high degree of accuracy. With FEM analysis, the computation time gets reduced when initial motor prototypes have to be analyzed for multiple iterations. FEM modeling ensures user friendly visualization of the models to spot the vulnerability and adaptability in the design. Magnets lose their necessary contribution in producing rotating magnetic flux if its thickness and pole embrace are taken too low. Being a 14-pole motor, upper limit of magnetic pole embrace is chosen in accordance with the rotor outer surface peripheral length for surface inset type PMSM, so that space between the magnets can be utilized for flux linkage and thereby reducing the flux leakage. An electric motor with

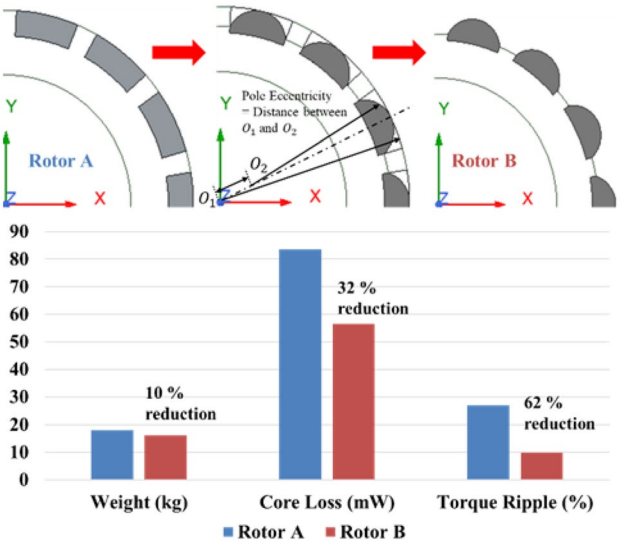


Fig. 16 Effect of rotor geometry

Table 8 FEM experiment results (Units for A and C parameters are in mm, B is unitless)

Experiment	*A*	*B*	*C*	Core Loss (mW)	
				Model I	Model II
1	3	0.65	150	178.7	123.4
2	3	0.75	150	175.5	119.5
3	3	0.65	160	190.6	131.6
4	3	0.85	160	183.7	123.0
5	3	0.75	170	199.0	135.4
6	3	0.85	170	195.2	130.7
7	4	0.65	150	167.9	108.2
8	4	0.75	150	165.3	104.7
9	4	0.65	160	179.1	115.4
10	4	0.85	160	172.9	108.2
11	4	0.75	170	187.3	118.7
12	4	0.85	170	183.7	114.9
13	5	0.65	150	159.3	96.9
14	5	0.75	150	156.7	93.9
15	5	0.65	160	169.9	103.4
16	5	0.85	160	164.4	97.1
17	5	0.75	170	177.6	106.4
18	5	0.85	170	174.6	103.2
19	6	0.65	150	152.0	87.7
20	6	0.75	150	149.9	85.4
21	6	0.65	160	162.2	93.6
22	6	0.85	160	157.2	88.1
23	6	0.75	170	169.9	96.8
24	6	0.85	170	167.1	93.6
25	7	0.65	150	146.1	80.3
26	7	0.75	150	144.0	78.3
27	7	0.65	160	155.8	85.7
28	7	0.85	160	151.1	81.1
29	7	0.75	170	163.2	88.7
30	7	0.85	170	160.6	86.2
31	8	0.65	150	137.2	74.2
32	8	0.75	150	138.7	72.3
33	8	0.65	160	149.8	79.2
34	8	0.85	160	145.7	75.2
35	8	0.75	170	157.2	82.0
36	8	0.85	170	154.8	79.9

small stack length produces lower torque during high-speed operation. Even though higher stack length results in higher torque, the torque reduces swiftly with increase in speed. Hence, stack length has to be optimum under considered constraints. Constraints for three design parameters are: 3 mm $\leq A \leq$ 8 mm, 0.65 $\leq B \leq$ 0.85, 150 mm $\leq C \leq$ 170 mm. Lower limits and upper limits of all three parameters are considered to obtain optimum torque with reduced weight and least core loss.

Table 9 Individual and associative coefficients of RSM

	y_1	y_2
β0	−16.0281	400.4877
β1	−11.0124	−37.9300
β2	39.5326	247.6950
β3	1.7875	−4.0040
β12	3.5295	−13.3768
β13	−0.0407	0.1171
β23	−0.3057	−3.3156
β11	0.6420	1.8694
β22	−24.0917	203.5722
β33	−0.0009	0.0208

The magnetic pole eccentricity is chosen as 2 mm to mitigate the torque ripple [33], and thickness of magnet is measured from the center of the pole periphery. Selection of rotor geometry provides the advantage of decrease in motor weight and core loss along with the reduction in torque ripples. For proving this, Model II with optimized geometry is analyzed with two designs of rotors: Rotor A having pole eccentricity equal to 0 mm and Rotor B having pole eccentricity equal to 2 mm. Figure 16 illustrates the comparison and percentage of reduction in core loss, torque ripple and weight of the motor by employing the proposed design of the rotor in SI-PMSM. The various FEM experiments and the corresponding core losses are tabulated in Table 8.

For implementing surrogate modeling, the response of interest is considered as a polynomial function $f(p)$ which includes all associative terms as given as follows:

$$f(p_1,p_2,p_3) = \beta_0 + \sum_{k=1}^{3} \beta_k p_k + \sum_{k=1}^{3} \sum_{l=k+1}^{3} \beta_{kl} p_k p_l + \sum_{k=1}^{3} \beta_{kk} p_k^2 \tag{23}$$

Here, terms β_0, β_k are individual coefficients and β_{kl}, β_{kk} are associative coefficients. These coefficients are computed by regression method with the help of observations of the output response. The 36 FEM experiments results are exploited for computation of coefficient matrix with respect to set of variables $[\{p_1, p_2, p_3\}_j]_{1:n}$ which is defined as design of experiments (DOE). Coefficients of RSM are calculated using MATLAB and are shown in Table 9.

Acknowledgements This research was supported by Project and Research Laboratory of Electrical Engineering Department in Delhi Technological University, New Delhi, India.

References

1. Baojun, Q.; Qingxin, Y.; Li, Y.; Miguel Angel, S.; Ma, S.: Li Z (2020) A novel surface inset permanent magnet synchronous motor for electric vehicles. Symmetry, MDPI **12**, 179 (2020)

2. Kolano, R.; Burian, A.K.; Krykowski, K.; Hreczka, M.; Polak, M.: Amorphous soft magnetic core for stator of the high speed PMBLDC motor with half open slots. IEEE Trans. Magn. **52**(006), 2016 (2016)
3. Zhu, X.; Cheng, M.; Xiao, F.; Sun, Y.: Comparison of doubly salient permanent magnet machines with E-shaped and π-shaped stator iron core segments. In: IEEE International Magnetics Conference (2015)
4. Feng, Y.; Zhang, C.: Core loss analysis of interior permanent magnet synchronous machines under SVPWM excitation with considering saturation. Energies **10**, 005–006 (2017)
5. Koo, M.-M.; Kim, J.-M.; Kim, C.-W.; Ahn, J.-H.; Hong, K.; Jang-Young, C.: Core loss analysis of permanent magnet synchronous generator with slotless Stator. IEEE Trans. Appl. Supercond. **28**(3), 1–4 (2018)
6. Guo, Y.; Jianguo, Z.; Lu, H.; Lin, Z.; Li, Y.: Core loss calculation for soft magnetic composite electrical machines. IEEE Trans. Magn. **48**(11), 3112–3114 (2012)
7. Yao, A.; Odawara, S.; Sugimoto, T.; Fujisaki, K.: Core loss properties of a motor with Nano crystalline Rotor and Stator cores under inverter excitation. IEEE Trans. Magn. **54**(11), 1–5 (2018)
8. Okamoto, S.; Denis, N.; Kato, Y.; Ieki, M.; Fujisaki, K.: Core loss reduction of an interior PMSM using amorphous stator core. IEEE Trans. Ind. Appl. **52**(3), 2261–2266 (2016)
9. Mohamed, E.E.: Electromagnetic performance of partitioned stator PM machines with dual C-core/E-core topologies. In: International Middle East Power System Conference (MEPCON) (2017)
10. Huang, Y.; Dong, J.; Zhu, J.; Guo, Y.: Core loss modeling for permanent magnet motor based on flux variation locus and finite element method. IEEE Trans. Magn. **48**, 1023–1026 (2012)
11. Ibrahim, M.; Pillay, P.: Advanced testing and modeling of magnetic materials including a new method of core loss separation for electrical machines. IEEE Trans. Ind. Appl. **48**, 1507–1514 (2012)
12. Gordon, R.S.; Liu, X.: Core losses in Permanent magnet motors. IEEE Trans. Magn. **26**, 1653–1655 (1990)
13. Russell, E.; James, K.: A New Optimizer using particle swarm theory. In: Sixth International Symposium on Micro Machines and Human Science, IEEE, pp. 39–43 (1995)
14. Marco, D.; Vittorio, M.; Alberto, C.: Ant system: optimization by a colony of cooperating agents. IEEE Trans. Syst. Man Cybern. **26**, 29–40 (1996)
15. Dervis, K.; Bahriye, B.: A powerful and efficient algorithm for Numerical Function optimization : artificial bee colony (ABC) algorithm. J. Global Optim. **39**, 59–471 (2007)
16. John, J.: Genetic algorithms. Sci. Am. **267**, 66–73 (1992)
17. Esmat, R.; Hossein, N.; Saeid, S.: GSA: a gravitational search algorithm. Inf. Sci. **179**, 2232–2248 (2009)
18. Seyedali, M.; Seyed, M.M.; Abdolreza, H.: Multi-verse optimizer: a nature inspired algorithm for global optimization. Neural Comput. Appl. **27**, 495–513 (2015)
19. Meeta, K.; Anand, J.H.; Suresh, C.: Socio evolution & learning optimization algorithm: a socio-inspired optimization methodology. Futur. Gener. Comput. Syst. **81**, 252–272 (2017)
20. Laith, A.; Daibat, A.; Mirjalili, S.; Mohamed, A.; Amir, H.: The arithmetic optimization algorithm. Comput. Methods Appl. Mech. Eng. **376**, 1–38 (2021)
21. Laith, A.; Daibat, A.: A Novel Hybrid Antlion optimization algorithm for multi-objective task scheduling problems in cloud computing environments. Clust. Comput. **24**, 205–233 (2020)
22. Laith, A.: Group Search Optimizer: a nature-inspired meta-heuristic optimization algorithm with its results, variants and applications. Neural Comput. Appl. **33**, 1–24 (2020)
23. Laith, A.; Dalia, Y.; Abd, E.; Ahmed, A.E.; Mohammed, A.; Amir, H.G.: Aquila Optimizer: a novel meta heuristic optimization algorithm. Comput. Ind. Eng. **157**, 1–59 (2021)
24. Shahrzad, S.; Seyedali, M.; Andrew, L.: Grasshopper optimization algorithm: theory and application. Adv. Eng. Softw. **105**, 30–47 (2017)
25. Jinlin, G.; Frederic, G.; Julien, T.C.; Yanliang, X.: Proposal of a kriging output space mapping technique for electromagnetic design optimization. IEEE Trans. Magn. **53**, 2017 (2017)
26. Fujishima, Y.; Katsuta, T.; Wakao, T.: Design optimization of PMSM by response surface methodology. J. Appl. Phys. **91**, 8305–8308 (2002)
27. Laith, A.; Daibat, A.: A comprehensive survey of the grasshopper optimization algorithm: results, variants and applications. Neural Comput. Appl. **32**, 15533–15556 (2020)
28. Faiz, J.; Ebrahimi, B.M.; Bilal, A.; Hamid, A.: Finite element transient analysis of induction motors under mixed eccentricity fault. IEEE Trans. Magn. **44**(1), 66–74 (2008)
29. Strangas, E.K.; Abdullah, A.; Khalil, H.K.: Speed Control of Permanent Magnet Synchronous Motor with uncertain parameters and unknown disturbances. IEEE Trans. Control Syst. Technol. **2020**, 1–8 (2020)
30. Paavo, R.; Floran, M.; Deepak, S.; Anouar, B.: Effect of mechanical stress on excess loss of electrical steel sheets. IEEE Trans. Magn. (2015)
31. Miller, T.J.: Brushless Permanent Magnet and Reluctance Motor Drives. Oxford Science Publication, Oxford (1989)
32. Juan, C.; Pavlos, S.; Eduardo, C.; Rafael, E.: Core lamination selection for distribution transformers based on sensitivity analysis. Electr Eng. **95**, 33–42 (2013)
33. Zhao, J.; Li, B.; Liu, X.; Li, X.; Chen, Z.: Research on the torque and back emf performance of a high speed PMSM used for flywheel energy storage. Energies **8**, 2877–2880 (2015)

Int J Syst Assur Eng Manag
https://doi.org/10.1007/s13198-021-01235-3

ORIGINAL ARTICLE

Assessing the impact of negative WOM on diffusion process

Deepti Aggrawal[1] · Mohini Agarwal[2] · Rubina Mittal[3] · Adarsh Anand[4]

Received: 15 January 2021 / Revised: 16 May 2021 / Accepted: 2 August 2021
© The Society for Reliability Engineering, Quality and Operations Management (SREQOM), India and The Division of Operation and Maintenance, Lulea University of Technology, Sweden 2021

Abstract The diffusion process has been considered as the propagation of messages associated with new ideas that lead to innovations; be it products, processes, or technology. With the anticipation of the change in receptor behavior, this diffusion process tends to bring out the adoption of the innovation. Most of the literature on innovation diffusion modeling is based on market growth however, very less work is available that focuses on how a new product penetrates a market under the effect of attrition on its growth. The intended purpose here is to study the dynamic behind the growth of an innovative product. The impact that past adopters of an innovation exercise on potential adopters by convincing them to imitate them in their choice to accept/reject the advancement (communication impact, impersonation impact), assists in explaining the acceleration of the diffusion process. With this objective, we have formulated and investigated an innovation diffusion model to include both adoption and disadoption behavior. The proposed framework has been validated and empirically analyzed on three real sales data sets.

Keywords Attrition · Diffusion · Disadoption · Innovators · Imitators · Predictive modeling

Abbreviations

IDP	Innovation diffusion process
WOM	Word-of-mouth
CRM	Customer relationship management
CRA	Customer relationship approach
PLC	Product life cycle

1 Introduction

The circulation of ideas associated with new product/services which leads to innovation which is termed as diffusion process (Bass 1969; Singh et al. 2012; Parvin and Beruvides 2021). Despite the plethora of work in diffusion modeling (Bass 1969; Mahajan Muller Wind 2000; Singh et al. 2012) the literature which has been the main thrust of research on the growth of single-purchase durable goods has not examined customer attrition. Customer attrition and its counterpart retention have gained considerable attention from managers and researchers due to the relationship between an organization's retention rate and its long-term profit (Huefner and Hunt 2000). There is a considerable influence on a new product/service due to customer attrition. Beginning from the initial stage of penetration into a market, some customers switch to competitors or leave the

✉ Mohini Agarwal
magarwal@amity.edu

Deepti Aggrawal
deeptiaggrawal@dtu.ac.in

Rubina Mittal
drmittal.keshav@gmail.com

Adarsh Anand
adarsh.anand86@gmail.com

[1] USME, Delhi Technological University, Delhi 110095, India

[2] Amity School of Business, Amity University, Noida, UP 201301, India

[3] Department of Mathematics, Keshav Mahavidyalaya, University of Delhi, Delhi 110034, India

[4] Department of Operational Research, University of Delhi, Delhi 110007, India

Published online: 09 August 2021

product segment. Moreover, some customers influence others in their buying decisions i.e., the spreadness of experience of existing adopters in shaping the opinion of potential adopters. Customer advocacy can be a strong driving force in the adoption of the product/services in which WOM plays a significant role.

WOM communication can be defined as the informal or the person-to-person communication between the communicator (adopter/user) and a receiver (yet to adopt) regarding either the brand/organization or product/services (Buttle 1996). Sometimes, it is regarded as the exchange of information between individuals to persuade individuals to adopt the offerings. In 1983, Richins (Richins 1987) had defined WOM as an act of communication in which adopters share the experience (satisfactory/unsatisfactory) earned with at least one person (friend/acquittance/family member). These communications are used for evaluative purposes wherein consumers relate this information as positive or negative information either for recommending or preventing others from unsatisfactory experiences. The impact that past adopters of an innovation exercise on potential adopters by convincing them to imitate them in their choice to accept/reject the advancement (communication impact, impersonation impact), assists in explaining the acceleration of the diffusion process. Several studies have examined the growth of competitive markets focused on the competition for attracting customers (Kumar and Krishnan 2002). Libai et al. (2009) have combined the diffusion modeling with CRA for examining the impact of attrition on the growth in the service market.

Following are the goals behind this study:

- Understand the attrition aspect is used in the context of disadoption generated due to NWOM.
- Understand how the disadoption rate affects the growth of a new product/service category.
- Develop a framework to understand how the dynamics of customer attrition affect the growth of consumer durables and examine the consequences from management perspectives.

The paper has been chronologically arranged as follows: after the brief introduction about the diffusion process and word-of-mouth communication, the emphasis has been to discuss the detailed literature as given in Sect. 2. Following this, the mathematical framework to explain the impact of negative WOM on IDP has been modeled in Sect. 3. In Sect. 4, the analysis to validate the proposed models has been carried on three different data sets. After it, the managerial implication and discussion have been provided in Sect. 5 and 6, respectively. Lastly, the conclusion in Sect. 7 and a list of references have been presented.

2 Background

Considerable research has been directed towards the understanding of the diffusion process and the impact of WOM communication. We will be talking about an amalgamation of these two schools of perception and how the diffusion process is being influenced by the WOM. In the strategic planning of an organization, the diffusion process has always played a key role. The introduction of a new product has been the survival mantra for an organization, but the diffusion process predominates it. As the adoption process of innovation generates revenue thereby making the diffusion process an important stimulus. The noteworthy work by Bass (1969) has been the base of many research work carried to understand the diffusion process and is extensively used for forecasting purposes. There are several extensions to it namely Horsky and Simon (1978) and Lilien and Rao (1978).

According to Horsky and Simon (1978), advertising has been the straight route for circulating knowledge about the presence of the product/service. Lilien and Rao (1978) pointed out the importance of individual marketing as the direct influences on the adoption. Dodson and Muller (1978) utilize the modified exponential curve to display the development of the product and the logistic pattern to show the expansion in the count of adopters. Easingwood et al. (1983) recommended that the fluctuating nature of WOM communication can be consolidated in the Bass model by considering the coefficient of internal influence varying over time.

Parker and Gatignon (1994) worked on developing the function form and understanding the impact of marketing mix variable on the diffusion process especially for existing or newly introduced brands in mature categories. Putsis Jr et al. (1997) addressed the issue of interaction patterns among the adopters in one country affecting the adoption pattern in other countries. Roberts and Lattin (2000) surveyed diffusion models based on an understanding of the dynamics at the level of the individual consumers. Their work has extremes: market level models in which each member has the same chance of adoption and the individual level model in which the chances of adoption by individuals have been considered different. Jain et al. (1991) focused on modeling the parsimonious diffusion model that can cater to both the dynamics of demand and supply-side restrictions. In the work by Kumar and Krishnan (2002) the diffusion process is developed to be flexible for accommodating the simultaneous impact of interactions between the lag and lead country. Ackerberg (2003) presented a dynamic learning model based on both 'informative', 'prestige', and 'image' effects of advertising. Empirical analysis to understand the different effects

of advertising for nondurable experienced market. De Bruyn and Lilien (2008) presented multi-stage dynamic models as components to consider and refine the comprehension of spontaneous, electronic references. In 2012, Kapur et al. (2012) built up a model based on the sales of the software to uncover the benefits and discover advertising strategies.

The basic Innovation diffusion process has been extended to understand the case of generic products like Agarwal et al. (2019) worked on developing a framework for the sales of generational products instilling awareness and adoption aspects. Sachdeva (2017), worked on the customer attrition aspect wherein the disadoption from innovators and imitators has been explicitly modeled in the form of the mean value function. In 2019, yet another work by Agarwal et al. (2017) focused on the dynamic behavior of market size and the concept of market coverage to understand the IDP. The literature on diffusion modeling is mainly concentrated on the growth pattern of consumer durable goods and very least importance on the aspect of customer attrition. Few researchers like Libai et al. (2009), Sachdeva (2017) and Anand et al. (2016) have explicitly examined customer attrition in services.

Several researchers have talked about the significant role of interpersonal communication after the contribution by Katz and Lazarsfeld (1955), in affecting the adoption patterns for the new product viz. Arndt (1967) and Sheth (1971). As quoted by Rogers (1962) in their work that impersonal sources are important at the awareness stage whereas personal sources have more influence in the evaluation stage of the adoption process. Swan and Oliver (1989) claimed that interaction among the consumers impacts their future purchases and behaviors. Henricks (1998), advocated that information from informal WOM combination is thought to strongly impact potential adopter's evaluation pattern about the products, organization, or any purchase decision in the future. In the last two decades, many studies came up to explain the categorization in negative WOM communication. It was Huefner and Hunt (2000) who classified the negative WOM as either a type of retaliatory action or an aggressive complaint behavior to adversely impact the business and secondly, to simply alert others from the associated risk. However, Sundaram et al. (1998) talked about the gist of negative WOM communication and divided it into four dimensions: 'altruism', 'anxiety reduction', 'vengeance', and 'advice-seeking'.

Researchers have claimed that the impact of negative WOM can have a more critical role in consumer's decision-making processes when managing intangibles than with tangible (File et al. 1992; Anderson 1998; Libai et al. 2009). Libai et al. (2009) have combined diffusion modeling with a customer relationship approach to investigate the influence of attrition on growth in service markets. Mesak et al. (2011) developed an optimization model for developing advertising policies and related the outcomes to the diffusion process in which the advertising affects the coefficient of innovation which is based on the disadoption rate. Hu (2015), researched how individuals' assessment rule changes, and which elements impact individuals' apparent behavior.

Seeing the abundance of literature in modeling and understanding the consumer adoption patterns based on several aspects of the market environment. This paper focuses on contributing to literature in examining the impact of negative WOM on IDP. The obtained mathematical form is in line with the equation of first-time purchasers as given by Bass (1969).

3 Modeling framework

Consumer behavior is dynamic, their expectation, and needs from specific products/services vary. Understanding the behavior via modeling their behavior pattern is the need of the hour. Few researchers have worked on the concept of attrition in the innovation diffusion process, but this field requires to be worked upon. The basic notations and assumptions used are as follows:

Notation

$N(t)$	is the cumulative number of adopters till time 't'.
m	is the potential market size.
p and q	are the coefficient of external and internal influence.
λ *and* b	is the rate of adoption.
β	is the learning parameter.
η	is the disadoption rate.

Assumption

1. The population of potential adopters who will adopt the innovation is constant.
2. Consumers are permitted to adopt one unit.
3. Geographical frontiers do not alter consumer behavior
4. The innovation propagation in the marketplace is in isolation.

To begin with, the mathematical formulation of the adoption behavior as given by Bass in 1969 (Bass 1969) as follows:

$$N(t) = m\left[\frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p}e^{-(p+q)t}}\right] \tag{1}$$

which is based on the concept that the innovators with word-of-mouth communication do propagate and spread the information and shape the opinion of non-adopters. In 2004, Kapur et al. (2004) gave an alternative notion to

understand the behavior of first-time purchasers by removing the barrier of people adopting through direct or indirect influence. The sales rate equation can be given as:

$$N(t) = m\left[\frac{1 - e^{-bt}}{1 + \beta e^{-bt}}\right] \tag{2}$$

The assumption that the characteristics of innovation or its perception do not change was relaxed in the model by Kapur et al. (2004). To understand the contagion effect of innovation in influencing the receptor behavior, the focus has been on understanding the impact of negative WOM on the total number of adoptions. The sales rate equation for the adoption process on the given set of assumptions can be given as follows:

$$\frac{dN(t)}{dt} = \lambda(t)[m - N(t)] - \eta p(t)[m - N(t)] \tag{3}$$

Equation (1) describes the sales intensity function for describing the impact of attrition on basic IDP as the intensity of first-time purchasers minus the intensity of disadoption due to the spread of negative WOM. To account for the decreasing probability of disadoption which decreases with time as the remaining number of potential purchasers starts decreasing. The following functional form $p(t)$ has been assumed:

$$p(t) = \frac{[m - N(t)]}{m} \tag{4}$$

The product $\eta p(t)$ as mentioned in Eq. (3) gives the instantaneous rate of disadoption. This rate is maximum at the beginning (equal to η) and tends to its minimum value (equal to 0) when the entire market is catered.

3.1 A new model with disadoption (proposed model I)

As mentioned above, $\lambda(t)$ accounts for the rate of diffusion for new product into the market. To simplify the constant functional form has been used i.e., $\lambda(t) = \lambda$. Solving the differential Eq. (3) together with (4) under the impact of negative WOM and assuming the initial condition as $N(t = 0) = 0$ the following equation can be obtained:

$$N(t) = m\left[\frac{(\lambda - \eta)\left(1 - e^{-\lambda t}\right)}{(\lambda - \eta) + \eta e^{-\lambda t}}\right] \tag{5}$$

On further simplification, Eq. (5) can be rewritten as

$$N(t) = m\left[\frac{1 - e^{-\lambda t}}{1 + \frac{\eta}{(\lambda - \eta)} e^{-\lambda t}}\right] \tag{6}$$

which is analogous to the Bass model (Bass 1969) as given by Eq. (1) and its alternative form Kapur et al. (1994) given by Eq. (2).

3.2 Proposed model II

In the alternative formulation to the Bass model (Bass 1969) given by Kapur et al. (2004), it was highlighted that the adoption process can be well described by considering the logistic form for the diffusion process. Thus, to supplement the same here the rate of diffusion for a new product has been assumed to be a logistic pattern i.e., $\lambda(t) = \frac{\lambda}{1+\beta e^{-\lambda t}}$. Using the differential Eq. (3) we have:

$$\frac{dN(t)}{dt} = \frac{\lambda}{1 + \beta e^{-\lambda t}}[m - N(t)] - \eta p(t)[m - N(t)] \tag{7}$$

Using Eq. (4) together with (7) and solving the differential equation under the impact of negative WOM and assuming the initial condition as $N(t = 0) = 0$ the following equation can be obtained:

$$N(t) = m + \frac{e^{-\lambda t}}{1 + \beta e^{-\lambda t}}\left[\frac{n}{m\lambda\beta}\ln\left(\frac{1 + \beta e^{-\lambda t}}{1 + \beta}\right) + \frac{1}{(1 - \beta)m}\right]^{-1} \tag{8}$$

Accordingly, the IDP under the impact of negative WOM as obtained is given by Eqs. (5) and (8).

4 Numerical illustration

To check the importance and validation of the model, analysis for the proposed model has been carried on three data sets. The first data set (DS I) relates to sales of Nokia mobile phone (Anand et al. 2014), the second data consists of sales for Samsung Smartphone and the third data pertains to sales for Acer Personal Computer (PC) (in millions) (Agarwal et al. 2019). For the parameter estimation purpose, the statistical software SAS Institute Inc (2004) has been used.

The estimated values of both the proposed models and the Bass model (Bass 1969) obtained for three different data sets are given in Tables 1, 2, 3, and 4. By looking for the disadoption rate for the two proposed models on DS I (Tables 1, 2), a higher value of adoption is obtained for the Proposed Model I (0.071637) compared to Proposed Model II (0.000772). Comparing for DS II (Table 3), a minor difference has been seen in the disadoption rate (0.0117) whereas for DS III (Table 4), a high value of disadoption rate has been obtained Proposed Model II in contrast to Proposed Model II with the difference of 0.131. Here it should be noted that results obtained for proposed model-I and the Bass model (Bass 1969) are in line. Thus, depicting that Bass model (Bass 1969) way back in 1969 was able to cater to both the positive as well as negative WOM though

Table 1 Comparison between present and prior research

	Present Research	Libai et al. (2009) and Sachdeva (2017)	Research on IDP*
Modeling the diffusion process of an innovation	Yes	Yes	Yes
Fluctuating nature of WOM communication	Yes	Yes	No
Modeling the impact of attrition on IDP	Yes	No	Yes
Examining the impact on consumer durables	Yes	No	Yes

*Research like Kapur et al. (2012), Ackerberg (2003), Dodson and Muller (1978).

Table 2 Parameter ESTIMATES for DS I (Anand et al. 2014)

Parameters	Proposed model I	Proposed model II	Bass model (1969)
m	1,036,232	4,720,286	1,036,232
λ	0.071637	0.000772	0.071637
η	0.061003	0.0042	–
β	–	1170	5.736529

Table 3 Parameter estimates for DS II (Agarwal et al. 2019)

Parameters	Proposed model I	Proposed model II	Bass model (1969)
m	2189.654	2169.235	2189.654
λ	0.163102	0.174795	0.163102
η	0.15815	0.000052	–
β	–	54.97032	31.93687

Table 4 Parameter estimates for DS III (Agarwal et al. 2019)

Parameters	Proposed model I	Proposed model II	Bass model (1969)
m	309.4476	264.9631	309.4476
λ	0.068836	0.1998	0.068836
η	0.032083	0.000275	–
β	–	68.3398	0.872911

it was not explicitly described in the research work. All the comparison criterion values are shown in Tables 5, 6 and 7.

The comparison criteria values as obtained in Tables 5, 6, and 7 shows the prediction capabilities of the proposed models. The value R^2 as obtained here for all three data sets is significantly close to one, depicting the close approximation to real data points and perfect fit to the data under consideration. On assessing Table 5, 6, and 7, it can be inferred that for DS I, the Proposed Model I and Bass

Table 5 Comparison criteria for DS I (Anand et al. 2014)

Criterion	Proposed model I	Proposed model II	Bass model
SSE	7.88E + 08	1.05E + 10	7.88E + 08
MSE	41,453,944	5.26E + 08	41,453,944
Root MSE	6438.5	22,925.8	6438.5
R^2	0.9971	0.9607	0.9971

Table 6 Comparison criteria for DS II (Agarwal et al. 2019)

Criterion	Proposed model I	Proposed model II	Bass model
SSE	34,584.9	192,926	34,584.9
MSE	1192.6	6652.6	1192.6
Root MSE	34.5338	81.5636	34.5338
R^2	0.9973	0.9848	0.9973

Table 7 Comparison criteria for DS III (Agarwal et al. 2019)

Criterion	Proposed model I	Proposed model II	Bass model
SSE	266.5	1993.7	266.5
MSE	8.0752	62.3025	8.0752
Root MSE	2.8417	7.8932	2.8417
R^2	0.9988	0.9907	0.9988

Model (Bass 1969) has performed equally well, whereas the same can be visualized for DS II and DS III.

Figures 1, 2, and 3 represent the goodness of fit curve for all the models for DS-I, DS-II, and DS-III respectively. Proposed Model II depicts the deviation from the Actual Sales, while Proposed Model I and Bass Model (Bass 1969) have shown quite a good superimpose on the Actual Sales values. The graphical representation between the actual data points and the predicted values depicts a close resembles among them and attributing to the fact the negative WOM does impact the sales. Thus, making it noteworthy to understand the diffusion process for a product/service by considering the impact of WOM communication in shaping the opinion of people who are yet to adopt the product.

5 Managerial implication

The present study depicts not only academic but also practical contributions in the field of modeling the diffusion process. Negative WOM communication is regarded as a means of expressing dissatisfaction toward a product/service. To minimize the impact of invisible negative WOM communication in the marketplace, business operators should understand the underlying reasons that generally lead to consumers' negative WOM communication. Results of this study provided insights into the impact of negative WOM communication on the intention of consumers. It is always important for marketers to understand the spread of information by existing adopters as it directly has an impact on the sales which in turn affects the profit generation of the company. The aforesaid model describes the impact of negative WOM on the adoption process. The lesser the rate of negative WOM the more beneficial it is for firms. Also, consumers may be less likely to spread negative WOM communication if they have an opportunity to express their comments or dissatisfaction to management directly. Thus, companies should focus on developing such a forum and help in fixing customer grievances associated with the adoption of the product. This will help in understanding consumers' expectations from the product

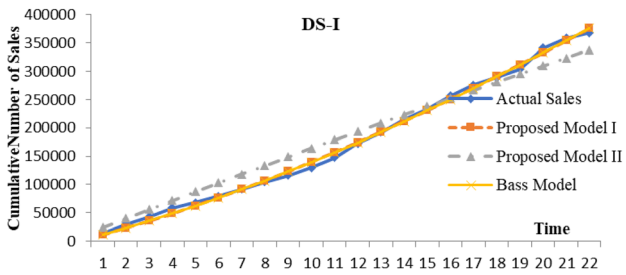


Fig. 1 Goodness of fit curve for DS-I

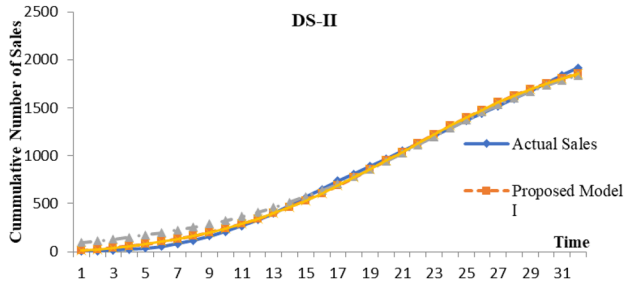


Fig. 2 Goodness of fit curve for DS-II

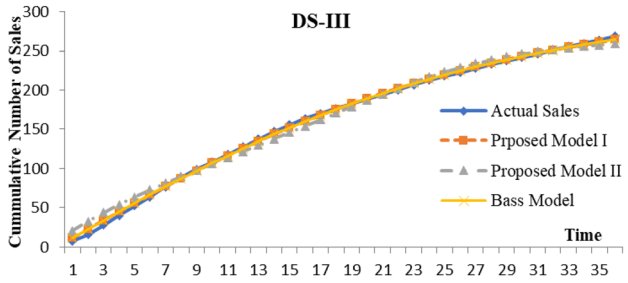


Fig. 3 Goodness of fit curve for DS-III

and what is the deviation between expectations and perception with the use of the product.

6 Discussion

The work presented here is to fill the research gap that exists in the literature about the modeling of the impact of the NWOM on the adoption process. The proposed models with their values of comparison criteria have depicted the importance of word of communication received by potential adopters in shaping their opinion. These features make the proposed model valuable to connect the weak link in the traditional adoption process. The proposed model is not to replace the traditional adoption process-based models but to enhance the research domain to include the mathematical aspects of NWOM. The proposed approach presents a new tool to assist management in understanding the entire adoption process and the role played by information sharing among the adopters and potential adopters. On the practical front, the approach is of high concern as the more the disadoption rate, the less will be the number of adopters. Thus, companies should strategically deal with and identify such dissatisfied adopters who might create the NWOM and consequently impact the sales. The work has only been analyzed for consumer durables which can be extended for other categories of the products. The present model has been developed for a single generation of the product which can be extended to the case of successive

generations with the differing rate of innovators and imitators.

7 Conclusion

Diffusion theory has always been an interesting and well-established field to discuss. Researchers have closely examined varied aspects of the growth of the product in the marketplace and how it has shaped the company's future. Few researchers (Gremler 1994; Elihu and Paul 1955; Lilien and Rao 1978; Sheth 1971; Sundaram et al. 1998; Swan and Oliver 1989; Dodson and Muller 1978) have also talked about the implications of WOM communication. The focus here was to understand and model the impact of negative WOM on the sales of the product. Through the present work, it was highlighted that the Bass model (Bass 1969) has talked about the adoption behavior due to WOM influence and has already taken care of its negative consequence on the final adoption of product/services. The results obtained are significant in explaining the market scenario with better predictive capabilities.

Acknowledgements The authors are thankful to the anonymous referees of the journal for their extremely useful suggestions to improve the quality of the article.

Funding No funding.

Declarations

Conflicts of Interest The authors confirm that there is no conflict of interest to declare for this publication.

References

Ackerberg DA (2003) Advertising, learning, and consumer choice in experience good markets: an empirical examination. Int Econ Rev 44(3):007–1040

Agarwal M, Aggrawal D, Anand A, Singh O (2017) Modeling multi-generation innovation adoption based on conjoint effect of awareness process. Int J Math, Eng Manag Sci 2(2):74–84

Agarwal M, Anand A, Bansal G, Pathak JP (2019) Innovation diffusion process based on market coverage under dynamic environment. Nonlinear Stud 26(3):517–526

Aggarwal R, Singh O, Anand A, Kapur PK (2019) Modeling innovation adoption incorporating time lag between awareness and adoption process. Int J Syst Assur Eng Manag 10(1):83–90

Anand A, Singh O, Agarwal M, Aggarwal R (2014, October). Modeling adoption process based on awareness and motivation of consumers. In: Proceedings of 3rd international conference on reliability, infocom technologies and optimization (pp. 1–6). IEEE

Anand A, Aggarwal R, Singh O, Aggrawal D (2016) Understanding diffusion process in the context of product dis adoption. *Научно-технические ведомости Санкт-Петербург, ского государственного политехнического университета. Экономические науки*. 2 (240)

Anderson EW (1998) Customer satisfaction and word of mouth. J Serv Res 1(1):5–17

Arndt J (1967) Role of product-related conversations in the diffusion of a new product. J Mark Res 4:291–295

Bass FM (1969) A new product growth for model consumer durables. Manag Sci 15(5):215–227

Buttle F (1996) SERVQUAL: review, critique, research agenda. Eur J Mark 30(1):8–32

De Bruyn A, Lilien GL (2008) A multi-stage model of word-of-mouth influence through viral marketing. Int J Res Mark 25(3):151–163

Dodson JA Jr, Muller E (1978) Models of new product diffusion through advertising and word-of-mouth. Manag Sci 24(15):1568–1578

Easingwood CJ, Mahajan V, Muller E (1983) A nonuniform influence innovation diffusion model of new product acceptance. Mark Sci 2(3):273–295

Elihu K, Paul L (1955) Personal influence. The Free Press, Glencoe

File KM, Judd BB, Prince RA (1992) Interactive marketing: the influence of participation on positive word-of-mouth and referrals. J Serv Mark 6(4):5–14

Gremler D (1994) Word-of-mouth communication: causes and consequences. Mark Rev 15(1):3

Henricks M (1998) Spread the word. Entrepreneur 26(2):120–125

Horsky D, Simon LS (1978, May) Advertising in a model of new product diffusion. In: TIMS/ORSA national meeting in New York City

Hu X (2015) Assessing source credibility on social media: an electronic word-of-mouth communication perspective (Doctoral dissertation, Bowling Green State University)

Huefner JC, Hunt HK (2000) Consumer retaliation as a response to dissatisfaction. J Consum Satisfaction, Dissatisfaction Complain Behav 13:61–82

Jain D, Mahajan V, Muller E (1991) Innovation diffusion in the presence of supply restrictions. Mark Sci 10(1):83–90

Kapur PK, Bardhan AK, Jha PC (2004) An alternative formulation to innovation diffusion model and its extensions. In: Kapoor VK (ed) Mathematics and information theory. Anamaya Publication, New Delhi, pp 17–23

Kapur PK, Pham H, Kumar V, Anand A (2012) Dynamic optimal control model for profit maximization of software product under the influence of promotional effort. J High Technol Managem Res 23(2):122–129

Kumar V, Krishnan TV (2002) Multinational diffusion models: an alternative framework. Mark Sci 21(3):318–330

Libai B, Muller E, Peres R (2009) The diffusion of services. J Mark Res 46(2):163–175

Lilien GL, Rao AG (1978) A marketing promotion model with word-of-mouth effect, Working Paper 976–78. Boston: Sloan School of Management, Massachusetts Institute of Technology

Mahajan V, Muller E, Wind Y (Eds.) (2000) New-product diffusion models (Vol. 11). Springer Science & Business Media

Mesak HI, Bari A, Babin BJ, Birou LM, Jurkus A (2011) Optimum advertising policy over time for subscriber service innovations in the presence of service cost learning and customers' disadoption. Eur J Oper Res 211(3):642–649

Parker P, Gatignon H (1994) Specifying competitive effects in diffusion models: an empirical analysis. Int J Res Mark 11(1):17–39

Parvin AJ, Beruvides MG (2021) Macro Patterns and trends of us consumer technological innovation diffusion rates. Systems 9(1):16

Putsis WP Jr, Balasubramanian S, Kaplan EH, Sen SK (1997) Mixing behavior in cross-country diffusion. Mark Sci 16(4):354–369

Richins ML (1987) A multivariate analysis of responses to dissatisfaction. J Acad Mark Sci 15(3):24–31

Roberts JH, Lattin JM (2000) Disaggregate-level diffusion models. New-product diffusion models. pp. 207–236

Rogers E (1962) Diffusion of Innovations. Free Press, New York

Sachdeva N (2017) Influence of customer attrition on diffusion of business education services. Int J Math, Eng Manag Sci 2(2):125–134

SAS Institute Inc (2004) SAS/ETS User's Guide Version 9.1, SAS Institute Inc., Cary, NC

Sheth JN (1971) Word-of-mouth in low-risk innovations. J Advert Res 11(3):15–18

Singh O, Anand A, Kapur PK, Aggrawal D (2012) Consumer behaviour-based innovation diffusion modelling using stochastic differential equation incorporating change in adoption rate. Int J Technol Mark 7(4):346–360

Sundaram DS, Mitra K, Webster C (1998) Word-of-mouth communications: a motivational analysis. Adv Consum Res 25(1):527–531

Swan JE, Oliver RL (1989) Post-purchase communications by consumers. J Retail 65(4):516–533

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

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

Attention U-Net Architecture for Dental Image Segmentation*

Pratyush Harsh
Department of Information Technology
Delhi Technological University
Delhi, India
pratyushharsh_2k17it88@dtu.ac.in

Ratul Chakraborty
Department of Information Technology
Delhi Technological University
Delhi, India
ratulchakraborty_2k17it95@dtu.ac.in

Shaurya Tripathi
Department of Information Technology
Delhi Technological University
Delhi, India
shauryatripathi_2k17it112@dtu.ac.in

Dr. Kapil Sharma
Department of Information Technology
Delhi Technological University
Delhi, India
kapil@dtu.ac.in

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498422

***Abstract*—Dental image segmentation is a very challenging task due to the presence of multiple teeth in the X-Ray and due to its different size in each human being. U-Net Architecture has been used in the Pancreas Segmentation [1] and Lung Segmentation [2] but the feature map of dental image is different from that of the lungs and pancreas. Technology used in the screening of the lungs and pancreas is CT Scan which is different form the standard X-Ray. To adapt the dental image feature we have proposed a U-Net Architecture [3] along with the attention gate [4]. U-Net Architecture [3] created by Olaf Ronneberger, Thomas Brox and Philipp Fischer is generally used in biomedical segmentation of images. It uses encoder and decoder architecture for the image segmentation but still there may be information losses while encoding the layers for feature map and unnecessary information is passed using the skip connection while decoding in standard U-Net Architecture. Limitation of the plain skip connection can be solved using the Attention block. Using Attention block in the U-Net model will further increase the sensitivity and prediction accuracy with minimum computational overhead. Model trained with attention block self-learn themselves. Any irrelevant regions in the input image are suppressed, and hence only the important features are highlighted. In our network architecture we have achieved the accuracy of 92.8 percent and 94.7 percent dice score.**

***Index Terms*—Attention, U-Net, Dental image segmentation, CNN.**

I. Introduction

X-ray images are an important aid to the dentist. They use these images to diagnose and treat patients. However, X-ray images are not easy to understand and even a professional may need years of experience. Teeth identification and teeth separation from the X-ray image attracts many in the field of medical science because of its application in the real world. Dental Image segmentation [5] is primarily used in the teeth classification. This is used to find the hidden abnormality in the dental structure like cavity, caries, etc. Using science and technology is the most interesting application of forensic identification. For example, forensic dentist relies on morphology for restoration to identify the victim who experienced the disaster. Biometric often gets lost after the postmortem or after a few days of the death of people but the dental structures of the person can be restored. In addition, with this other used for this is age estimation, plaque analysis etc. Dental Segmentation is in great interest.

II. Related Works

A. Fuzzy C-Means [6]

Fuzzy C-Means [6] is a clustering method allowing that data points can be in more than one cluster. In 1973, Dunn developed it, further improved in 1981 by Bezdek. It can be called a generalization of the K Means algorithm [7]. This method is very often used classifying pixel values of images with no distinction boundaries between their objects. However, it does not use any spatial or local information, making it prone to noises and image artifacts.

B. Global Threshold [8] [9]

This methodology revolves around choosing a threshold value used to separate the background and the object. The object pixels and background pixels are separated into two modes of gray levels. Having pixels of different intensity values makes the separation visible. Though, global thresholding hampers the accuracy of the predicted model, since factors such as ambient illumination, noise, gray level complexities of the image background and objects, makes the thresholding difficult.

C. Niblack Local Thresholding [10]

Niblack's algorithm [10] is a local thresholding technique which determines a threshold value to each image pixel over a gray level image. Let an image pixel be denoted by (p,q). Local threshold value is calculated inside a window size wxw by:

$$T_{(p,q)} = m_{(p,q)} + k * \delta_{(p,q)} \quad (1)$$

$m_{(p,q)}$ denotes the local mean and delta$_{(}$p,q) denotes the standard deviation of the image pixels inside the window. Bias

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:58:42 UTC from IEEE Xplore. Restrictions apply.

k is a constant value according to the contrast in neighboring pixels.

D. Sauvola Local Thresholding [10] [11]

Sauvola Local Thresholding is a modified version of Niblack Local Thresholding. For cases where the background contains light textures, Niblack's method does not work well as gray values start exceeding the threshold values. Local threshold value is calculated inside a window size wxw by.

$$T_{\text{Sauvola}} = m * (1 - c * (1 - s/R)) \tag{2}$$

Here, c is stated as the control factor between 0.2 and 0.5. Maximum standard deviation is stated by R (upto 128 for a grayscale doc).

E. Squeeze and Excitation Network [12]

This is an architectural unit which can be plugged into any convolutional network to improve the quality of the convolutional network by partially increasing the parameters. This block explicitly models channel relationship and channel interdependencies, it also includes a form of self-attention on channel. Squeeze and excitation (SE) block is hence used to improve accuracy of convolutional neural layers.

$$s = F_{\text{ex}}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \tag{3}$$

δ stated the ReLU function. For the gating function, Sigmoid is used. Excitation operations are conducted to identify channel-wise dependencies and learn non-linear/non-mutual relationships. After the global max pooling two fully connected layers for channel descriptor. In order to obtain the final block, transformations are rescaled with transformation output U.

III. PROPOSED NETWORK

A. Overview and Motivation

Concept behind U-Net is that we fuse the feature while encoding and decoding it in the decoder network. It gives data from encoder to decoder path and helps in recovering spatial information lost by max pooling operation. While information is passed through the skip connection unnecessary information is also passed and this can be reduced by adding an attention gate in the skip connection.

B. Network Architecture

Fig 2 displays network architecture. It consists of an Encoder part on left side, also known as a contracting part as well as expansion along with sandwich Attention block on the right side. Basic image segmentation uses this standard encoder and decoder network. It consists of multiple 3x3 convolution (unpadded convolution) each following a rectified

linear unit (ReLU). Also, max pooling operation of size 2x2 is used for down sampling. All the convolutional layers have been initialized with the normal initializer. While down sampling, every step leads to doubling the number of channels. Expansion path contains the copy layer and the Attention block

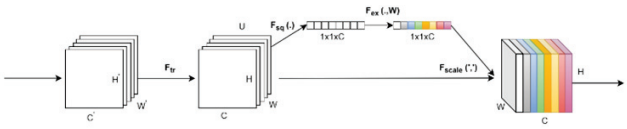


Fig. 1. Attention Block Capture the channel wise attention

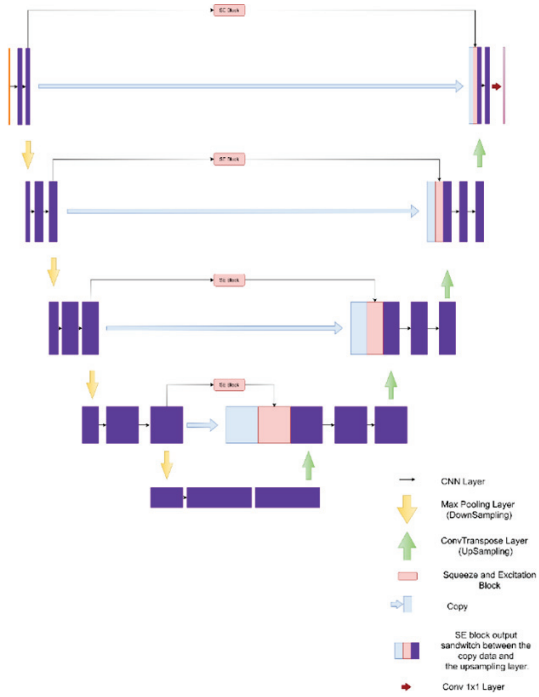


Fig. 2. Network Architecture

sandwich between the copy layer and the up-convolutional layer. While upscaling the number of channels at every up-convolutional step gets halved. Since the height and width dimension of the layer gets halved at every step, we need to choose the input size of the image according to that in our case we have chosen 512x512 as our input image. In the last layer of the network to match the output with the desired output the 16-channel using the 1x1 convolutional at last layer the channel is reduced to our desired output.

IV. PREPROCESSING AND TRAINING

A. Dataset

UFBA-UESC dental image dataset is used which is obtained from the Ivisionlab [13]. It contains 1500 dental x-ray images along with their manually segmented teeth data of 10 different categories. Images from all the folders is combined into two folder segmented image and ground truth. Division is done into the training, validation and test set in the ratio 70: 10: 20. Attention U-Net Architecture implemented with the Python and TensorFlow framework and experimented on NVIDIA RTX 2060.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:58:42 UTC from IEEE Xplore. Restrictions apply.

B. Preprocessing

Images from all the 10 folders are renamed and moved to new folder along with the segmented image. Input images are then resized to 512x512 pixels and then it is passed in the batch of 4 to the model

V. PERFORMANCE MEASURES

We have chosen the following set of metrics to evaluate our algorithm. Since the output image has only one channel and the pixels are either white or black. We can consider 0 and 1 for white and black regions.

A. Accuracy

Accuracy can be defined as the total number of predicted pixels that were correct out of all the pixels. More formally, it is defined using the formula

$$Accuracy = \frac{TrueNegative+TruePositive}{TruePositive+FalseNegative+FalsePositive+TrueNegative} \quad (4)$$

B. Precision

Precision is the ratio of correct predicted positive pixels to all the predicted positive pixels.

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \quad (5)$$

C. Recall

Recall can be defined as the ratio of correct predicted positive pixels to all the actual true pixels.

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \quad (6)$$

D. F1 Score

F1 Score is computed as twice the product of precision and recall divided by the sum of precision and recall.

$$F1Score = \frac{2 * (Precision * Recall)}{Precision + Recall} \quad (7)$$

E. Dice Score

Dice score is computed as two times the True Positive divided by the sum of False Positives, False Negatives and two times the True Positives

$$DiceScore = \frac{2*TruePositive}{2*TruePositive+FalsePositive+FalseNegative} \quad (8)$$

VI. EXPERIMENT RESULT

Table 1 displays various parameters utilized in this model. Rectified Linear Unit (ReLU) activation function, it is a linear function which is responsible for weighted input from a node towards output. Rectified Linear Unit function quicky converges training set with low computing power also better getting good performance metrics. Soft Dice Loss works better in class imbalance (some segments occupy less pixels/voxels than others) than the binary crossentropy. Adam Optimizer is used to quickly converge to the result.

Table 2 displays summary of the results. Attention U-Net Architecture has highest accuracy, F1 Score, Precision and Dice Score when contrasted with some other segmentation techniques. Fig 4 display the contrasting feature channels using attention network and Fig 4 displays the features channel without the attention gate. From Figure 5, we can depict that channels in attention block capture information from different regions, which can be further used for classification purpose. Fig 3 displays segmented images using other existing segmentation algorithm. Attention U-Net surpass all the other state of art algorithm for image segmentation.

Table 1

Model	Parameters
Activation Function	Rectified Linear Unit (ReLU)
Optimizer Function	Adam Optimizer
Loss Function	Soft Dice Loss
Batch Size	4
Total epochs	20
Validation Set Split	0.10
Attention Gate	SE-Block

CONCLUSION

Image segmentation is a preprocessing step in analyzing medical imaging as it deals with the X-Ray, CT-Scan of Lungs, Pancreas, Teeth, Brains and many more organs and segmentation which require a lot of skilled human effort. We have suggested a hybrid model integrating U-Net Architecture along with an attention gate that can capture and adapt the model for features of dental images (X-Ray), different from the features map of other organs (Lungs, Pancreas). Proposed model is also validated on standard segmentation technique. We have compared some metrics like Accuracy, Recall, Precision, F1 Score and Dice Score is used to validate proposed model. U-Net Attention architecture helped in outperforming existing segmentation algorithms with accuracy up to 92.8 percent as well as dice score up to 94.7 percent. Using the attention block in the U-Net Architecture self learns to suppress the irrelevant region and learns to highlight the salient feature for specific tasks.

ACKNOWLEDGMENT

We thank Dr. Kapil Sharma (Head of Department, Information Technology, Delhi Technological University) for being a constant support throughout the research. We also thank our colleagues from Delhi Technological University, who helped us with conclusive insights to make this research a success. We also thank Dr. Luciano Oliveira, Computer Science Department, UFBA for giving us access to UFBA UESC Dental Images Dataset for conducting our research. We would also like to express our gratitude towards our parents and closed ones, who provided us with constant moral and financial support.

REFERENCES

[1] E. K. a. R. Ceylan, "Pancreas Segmentation in Abdominal CT Images with U-Net Model," in 28th Signal Processing and Communications Applications Conference (SIU), Gaziantep, Turkey, 2020.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:58:42 UTC from IEEE Xplore. Restrictions apply.

Table 2

Algorithms	Parameters				
	Accuracy	*Precision*	*Recall*	*F1 Score*	*Dice Score*
Fuzzy C-Means	0.8235	0.6186	0.4525	0.4939	0.5328
Global Threshold	0.7929	0.5202	0.6931	0.5621	0.6243
Niblack Local Thresholding	0.8182	0.5129	0.8257	0.6138	0.6428
Sauvola Local Thresholding	0.8071	0.5028	0.8368	0.6282	0.6381
Attention U-Net	0.9280	0.9921	0.7921	0.8809	0.9470

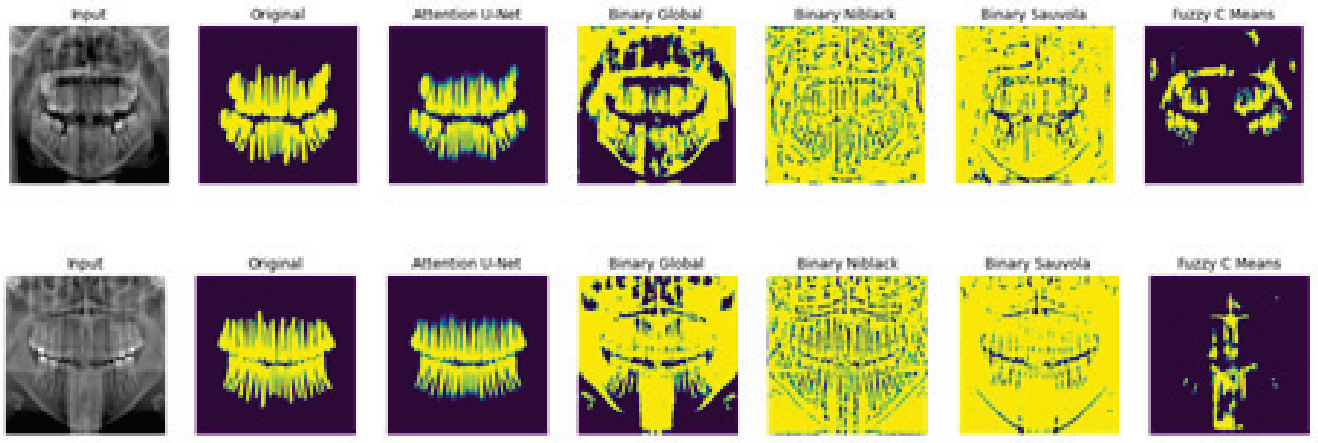


Fig. 3. Experiment Analysis of various segmented method

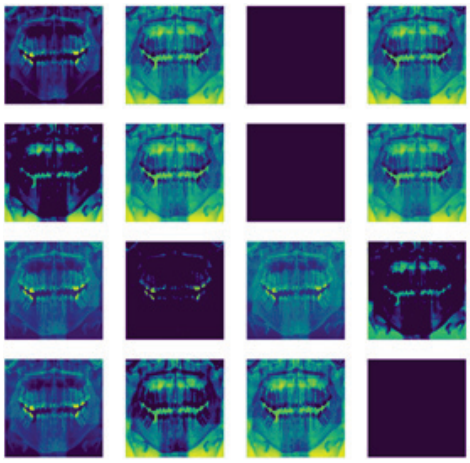


Fig. 4. Channel Feature map after the attention block

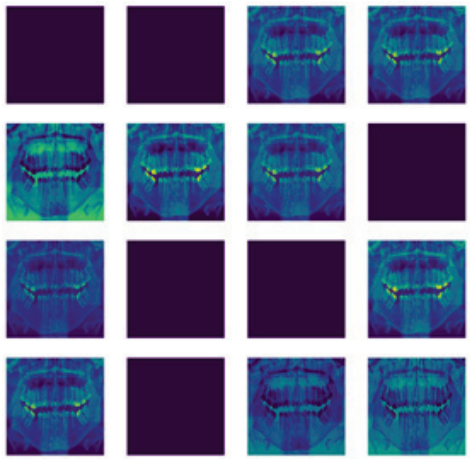


Fig. 5. Channel Feature map before attention block

[2] K. S. a. R. Z. H. Shaziya, "Automatic Lung Segmentation on Thoracic CT Scans Using U-Net Convolutional Network," in 2018 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2018.

[3] O. R. a. P. F. a. T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," 2015.

[4] O. O. a. J. S. a. L. L. F. a. M. L. a. M. H. a. K. M. a. K. M. a. S. M. a. N. Y. H. a. B. K. a. B. G. a. D. Rueckert, "Attention U-Net: Learning Where to Look for the Pancreas," 2018.

[5] J. F. M. R. K. A. M. P. a. L. O. G. Jader, "Deep Instance Segmentation of Teeth in Panoramic X-Ray Images," in 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Parana, 2018.

[6] M. K. Alsmadi, "A hybrid Fuzzy C-Means and Neutrosophic for jaw lesions segmentation," Ain Shams Engineering Journal, vol. 9, no. 4, pp. 697-706, 2018.

[7] E. K.-M. Algorithm, "Extended K-Means Algorithm," in 2013 5th International Conference on Intelligent Human-Machine Systems and Cybernetics, Hangzhou, 2013.

[8] H. Z. a. R. K. F. Yan, "A Multistage Adaptive Thresholding," Science Direct , 2005.

[9] A. A. D. N. a. N. J. R. Indraswari, "Teeth segmentationon dental panoramic Radiographs using decimation-free directional filter bank thresholding and multistage adaptive thresholding,," in International Conference on Information Communication, Technology and Systems (ICTS), 2015.

[10] W. Niblack, An introduction to digital image processing, Englewood Cliffs : Prentice-Hall, 1986.

[11] J. S. a. M. Pietikainen, "Adaptive Document Image Binarization," The Journal of the Pattern Recognition Society, vol. 33, pp. 225-236, 2000.

[12] L. S. a. G. S. J. Hu, "Squeeze-and-Excitation Network," in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, 2018.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:58:42 UTC from IEEE Xplore. Restrictions apply.

[13] T. Tuan, "Dental segmentation from X-ray images using semisupervised," Engineering Applications of Artificial Intelligence, vol. 59, p. 186–195, 2017.
[14] P. C. G. S. R. K. a. S. R. M. S. Sivagami, "UNet Architecture Based Dental Panoramic Image Segmentation," in 2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), Chennai, 2020.
[15] O. a. F. P. a. B. T. Ronneberger, "U-Net: Convolutional Networks for Biomedical Image Segmentation," Cham, 2015.
[16] O. O. M. S. M. H. B. K. B. G. D. R. Jo Schlemper, "Attention gated networks: Learning to leverage salient regions in medical images," Science Direct, pp. 197-207, 2019.
[17] X. L. L. X. a. Y. G. J. Wang, "A Diversified Supervised based U-shape Colorectal Lesion Segmentor with Meaningful Feature Supplement and Multi-Level Residual Attention Mechanism," in IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Seoul, Korea (South), 2020.
[18] A. O. P. M. M. V. a. T. S. M. Chauhan, "Blockchain and scalability," in In 2018 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C), Lisbon, 2018.
[19] C.H. Huang and C.Y. Hsu, "Computer-assisted orientation of dental periapical radiographs to the occlusal plane," Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology, vol. 105, Medicine, Oral Pathology, Oral Radiology, and Endodontology, vol. 105, no. 5," in Computer-assisted orientation of dental periapical radiographs to the occlusal plane, 2008, p. 649–653.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:58:42 UTC from IEEE Xplore. Restrictions apply.

2021 Fourth International Conference on Computational Intelligence and Communication Technologies (CCICT) | 978-1-6654-2392-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/CCICT53244.2021.00035

Comparative analysis of acquisition methods in digital forensics

Himanshu[1]
Department of CSE
AIACT&R, GGSIPU University
Delhi, India
himanshudubey869@gmail.com

Shobha Bhatt[2]
Department of CSE
Netaji Subhash University of Technology, East Campus
Delhi, India
bhattsho@gmail.com

Geetanjali Garg[3]
Department of CSE
Delhi technological University
Delhi, India
Geeta.anjali.garg@gmail.com

Abstract: This paper presents a comparative analysis of acquisition methods in digital forensic. The author compares the data acquisition and memory acquisition with the help of the Forensic toolkit imager (FTK imager) and ProDiscover tool is open source digital forensic tools. By taking some perimeter for comparison that is generally related to the data acquisition and memory acquisition. In this paper, data acquisition of a 32GB Pen drive is made with the help of both the tools, and the result of the data is taken and captured, which is also presented in this paper with the help of screenshot, and memory acquisition of physical memory (4GB RAM) is also made, and the result is captured and presented with the help of screenshot in this paper. By comparing acquisition methods with the help of these two tools, the author gets that the one tool is better than the other on certain perimeters and the other tool is also better than the first tool in some manner, which will help in the future to the researchers while they are using these two tools.

Keywords: Digital forensics, Data acquisition, Memory acquisition, Acquisition, Pen drive, Image, Digital evidence, Pro Discover, FTK imager

1. Introduction:

Acquisition is the procedure of making the same duplicate of data occupying the digital evidence like RAM, Hard Drive, CD, Pen Drive, or any electronic storage device [1]. It is making a byte by byte copy of the evidence which is to be found in digital devices. It makes the exact image of the physical memory [2] and the data. And the image is examined by the digital investigator. So, this paper presents the two acquisition methods, which are Data acquisition [3] and memory acquisition. There are two methods of acquisition which are as follows.

The researchers have faced many challenges, like the hardware interface of flash drives is streamlining in nature which may cause an issue while creating the image of the devices. [4]. As encryption usage is increasing day by day, the encrypted data that are vanished could be recovered, but there is a little problem while attempting to decrypt that data. As the size of the storage is rising gradually, the time taken for the scanning and creating disk image is also increased. As the data is saved over several sections in the cloud, which are distributed in nature over the network, it is difficult to search the sections during analysis [5]. While making the data acquisition through the FTK imager of a CD, it requires the Fernico device for making the data acquisition; without the Fernico device, the data acquisition of a CD is not performed. And the memory acquisition with the help of pro discover is also quite difficult to do because there is an error that randomly pops up and disturbs the operation.

Data acquisition and memory acquisition are widely used in the industry and the digital investigation process. It is also used in data recording and nvh testing. In this paper, a comparative analysis of the acquisition method is done based on the two tools that are FTK imager and Pro Discover while taking eight parameters for comparison after implementing the acquisition process and result is taken; then the comparison is done onto those eight parameters.

Section 2 is about a literature review. Section 3 elaborates on the data acquisition and different types of devices from which the evidence is collected. Section 4 is all about memory acquisition and its block diagram. Section 5 explained the data acquisition using the FTK imager tool, and section 6 explained the memory acquisition using the FTK imager. Section 7 explained the data acquisition with the help of ProDiscover. And then, section 8 shows the comparative study among both the tools, and section 9 contains the conclusion of this paper.

978-1-6654-2392-2/21/$31.00 ©2021 IEEE
DOI 10.1109/CCICT53244.2021.00035

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:13:20 UTC from IEEE Xplore. Restrictions apply.

2. Literature review:

Varsha Karbhari Sanap and Vanita Mane 2015, the author elaborates on the comparative study and simulation on digital forensic tools. The authors also do the case study on the Winhex tool, Pro Discover and active file recovery. Certain steps are to be taken to solve that case study, and then the comparative analysis is done on the three tools. The authors also presented a three steps investigation process which includes three steps to perform the digital investigation. These steps are acquisition analysis and reporting [6].

Brian Carrier 2002, the author elaborates about the open-source digital forensic tools in this paper with taking the legal argument along with that. The author stated that the tool used in the analysis and the digital investigation process must be reliable and relevant in nature. The author checks the tools which fall under the government guidelines and also checks the closed source tools. After comparing author finds that the open-source digital forensic tools are clearly comprehensive with the government guidelines as compared to closed source tools. The author also presented the three-phase forensic analysis process, which contains the acquisition, analysis, and presentation phase [7].

Josiah Dykstra and Alan T. Sherman 2012, the authors elaborate on the digital investigation process and related aspects. They have also described the features of the tools and do a comparative analysis of the tools while taking certain parameters to make the comparison. There are certain tools that are running on different platforms. And certain tools have their own specific features. There are some basic rules like the collected data should not be changed anytime [8].

FTK imager is founded by the Access Data Group. Also, they have the service of certification of tool training. Over 131,292 judiciary firms and Government Organizations use FTK around the countries. It analyzes a hard Disk searching for the information. It has the feature of discovering the deleted image and search a drive for words to take them as a password vocabulary to breach encryption. The image of a hard drive in one segment will be reconstructed with the help of this tool. It computes the MD5 checksum and SHA1 checksum to check the integrity of the image which is created by the FTK imager. The image could be stored in various formats like DD/raw, AD1, and E01.

Pro Discover was launched in 2001 by the ARC group. Remote forensic Power was first implemented by Pro Discover [9]. This tool is used in 80 countries in several complicated cases and various high-profile cases. It also addresses a broad range of cyber-crime frameworks experienced by legal agencies and organizations' internal security examiners. It is broadly used in incident response, digital forensic, and computer forensic. It also has the feature of policy-based electronic discovery. It has the feature of memory acquisition and Data acquisition.

3. Data acquisition: The procedure of gathering digital media/evidence from digital devices that have to be later investigated is known as acquisition. It stores the position of a system so that it could be later examined. The purpose of this phase is to keep all the digital values. In this process, the assigned and unassigned hard drive areas are copied, generally known as images. In this phase, the data is to be copied from the suspicious storage device to an authorized device. Evidence can be collected from different devices such as Pen drives, Digital cameras, embedded devices and chipsets, Physical hard drives, Storage cards, Optical media.

4. Memory acquisition: the copying of the data from physical memory to the storage device for preservation is known as memory acquisition [10]. In this process, the data from volatile to non-volatile storage is copied. It basically involves three processes. Given below, Figure 1 shows the block diagram of memory acquisition. Firstly, the capturing of physical memory is done; then the dumping rules apply to that then after that, by using sampling, the perfect image of the physical memory is extracted.



Figure 1: Block diagram of memory acquisition

5. Implementation of Data acquisition by using the FTK imager:

Here data acquisition is to be done by using the FTK imager tool in which the author has a pen drive of 32GB, or we can say that if we have a suspected pen drive with the help of FTK imager, we create the image of the pen drive then after we

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:13:20 UTC from IEEE Xplore. Restrictions apply.

read that file then we can understand what data is deleted and what is present and what is modified in the pen drive along with the date and time is shown by the fig 3. So here, we can show you some analysis of the data with the help of a screenshot. FTK imager also verifies the image after creating it; it is shown in figure 2.

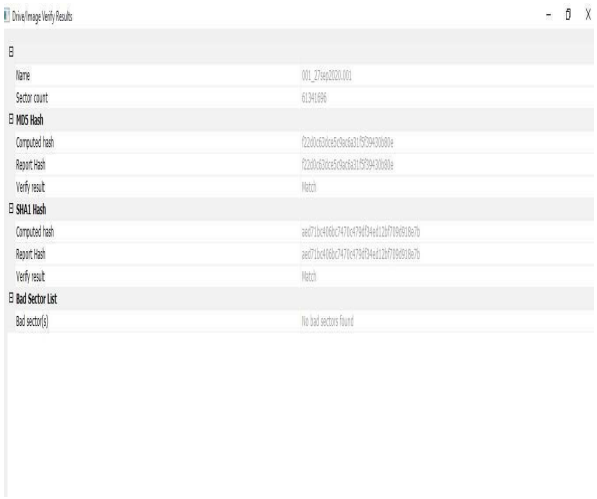


Figure 2: Verify image result

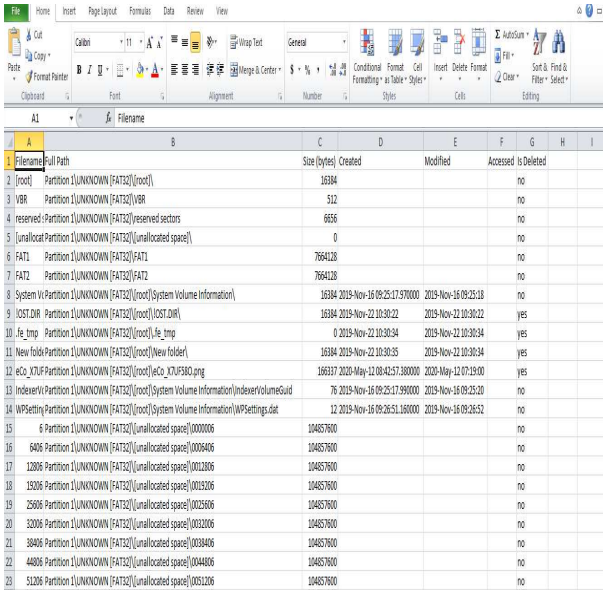


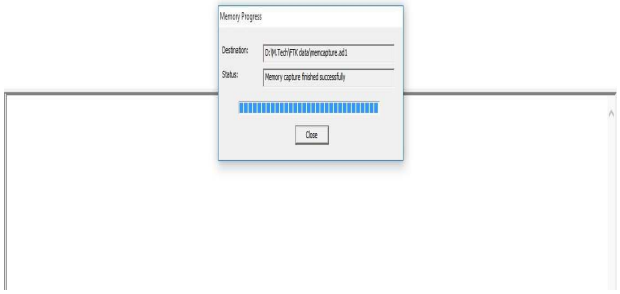
Figure 3: Image analysis

Information for D:\M.Tech\FTK data\001_27sep2020:
Physical Evidentiary Item (Source) Information:
[Device Info]
Source Type: Physical
[Drive Geometry]
Cylinders: 3,818
Tracks per Cylinder: 255
Sectors per Track: 63
Bytes per Sector: 512
Sector Count: 61,341,696
[Physical Drive Information]
Drive Model: SanDisk Cruzer Blade USB Device
Drive Serial Number: 4C530000150708101322
Drive Interface Type: USB
Removable drive: True
Source data size: 29952 MB
Sector count: 61341696
[Computed Hashes]
MD5 checksum: f22d0c63dce5c9ac6a31f5f39430b80e
SHA1 checksum: aed71bc406bc7470c479df34ed12bf709d918e7b
Image Information:
Acquisition started: Sun Sep 27 16:58:16 2020
Acquisition finished: Sun Sep 27 17:31:45 2020
Segment list:
D:\M.Tech\FTK data\001_27sep2020.001 to
D:\M.Tech\FTK data\001_27sep2020.020
Image Verification Results:
Verification started: Sun Sep 27 17:31:46 2020
Verification finished: Sun Sep 27 17:43:12 2020
MD5 checksum: f22d0c63dce5c9ac6a31f5f39430b80e: verified
SHA1 checksum: aed71bc406bc7470c479df34ed12bf709d918e7b : verified

6. Implementation of memory acquisition by using the FTK imager

Here memory acquisition is to be done by using the FTK imager tool in which author performs the memory acquisition of a laptop's physical memory (4GB RAM); fig 4 shows the capturing of memory [11] [12].





Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:13:20 UTC from IEEE Xplore. Restrictions apply.

Figure 4: Capturing memory

```
Information for D:\M.Tech\FTK data\memcapture.ad1:
[Computed Hashes]
 MD5 checksum:    58f633f858c145b415cafaac9e28ed4e
 SHA1 checksum:   883499da00c219beb89e7f83d7338fd1eb4ab1ed

Image information:
 Acquisition started:   Wed Nov 18 08:04:18 2020
 Acquisition finished:  Wed Nov 18 08:14:32 2020
 Segment list:
  D:\M.Tech\FTK data\memcapture.ad1
```

Figure 5: Analysis of memory image

7. Implementation of Data acquisition with the help of ProDiscover

Here, data acquisition of a 32GB pen drive with the help of the ProDiscover tool is made. After that author reads the file and understands what data is deleted and what is present and what is modified in the pen drive along with the date and time is shown by the fig 5, And fig 6 shows the result in which all the pen drive built-in related data and the data contained by the pen drive is also shown.

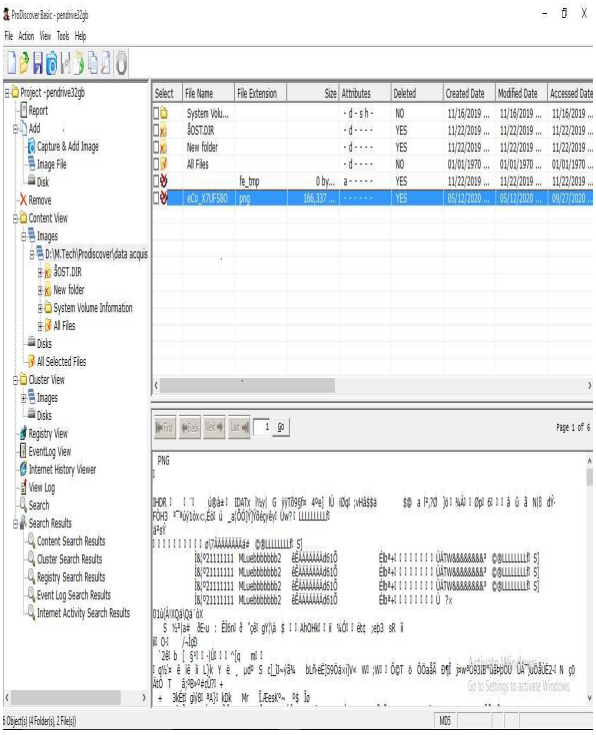


Figure 6: Image analysis

```
<IDate> 12/10/2020 </IDate>
<ITime> 11:01:21 </ITime>
<IChecksum> e8f42dabc9a9b779e0f3da719375ee6f </IChecksum>
<IChecksumValidated> 0 </IChecksumValidated>

<Drive>
        <DiskName> D:\M.Tech\Prodiscover\data acquisition.eve </DiskName>
        <OIVolName> NO NAME     </OIVolName>
        <OIVolSerNum> D80D-CDD6 </OIVolSerNum>
        <OIFileSystem> FAT32 </OIFileSystem>
        <OIBPS> 512 </OIBPS>
        <OITotClust> 1915991 </OITotClust>
        <OITotSec> 61341664 </OITotSec>
        <OIHiddenSec> 32 </OIHiddenSec>
        <OITotCapacity> 30670832 </OITotCapacity>
        <OIStartSec> 0 </OIStartSec>
        <OIEndSec> 61341663 </OIEndSec>

        <EvidenceOfInterest>
                <File>
                        <EOIFileName> D:\M.Tech\Prodiscover\data acquisition.eve\System Volume Information\
                        <EOIFCheckSum> MD5 Checksum: 41412EA12E152D4018D34E38296C74BE </EOIFCheckSum>
                        <EOICreated> 11/16/2019 09:25 </EOICreated>
                        <EOIModified> 11/16/2019 09:25 </EOIModified>
                        <EOIAccessed> 12/10/2020 00:00 </EOIAccessed>
                        <EOISTDInfoUpdated>  </EOISTDInfoUpdated>
                        <EOISTDINFOCreated>  </EOISTDINFOCreated>
                        <EOISTDINFOAccessed>  </EOISTDINFOAccessed>
                        <EOISTDINFOModified>  </EOISTDINFOModified>
                        <EOIMFTUpdated>  </EOIMFTUpdated>
                        <EOIMFTCreated>  </EOIMFTCreated>
                        <EOIMFTAccessed>  </EOIMFTAccessed>
                        <EOIMFTModified>  </EOIMFTModified>
                        <EOIFStartSec> StartSec: 4 </EOIFStartSec>
                        <EOIFADS>  </EOIFADS>
                        <EOIFSize>          76 bytes </EOIFSize>
                        <EOIBatesNum>  </EOIBatesNum>
```

Figure 7: Result

8. Comparative study:

Table 1: Comparative study of acquisition method by using the FTK imager and pro discover tool

Serial no.	Perimeter for comparison	FTK Imager	Pro discover
1.	Time is taken for data acquisition	33 minutes	37 minutes
2.	Accuracy of data acquisition	More accurate	Less accurate
3.	Image verification	Yes	No
4.	Progress bar estimation	No	Yes
5.	Checksum	Yes	Yes
6.	Sector count	61341696	61341664
7.	Recovery of file	Yes	Yes
8.	Time taken for memory acquisition	10 minutes	10 minutes 26 seconds

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:13:20 UTC from IEEE Xplore. Restrictions apply.

The table presents the comparison of acquisition methods with the help of FTK imager and pro discover tool. The table presents the comparison based on eight parameters according to their performance, as shown in this paper above. While comparing on the base of performance, it is found that the FTK imager is slightly better than the pro discover tool in both data acquisition and memory acquisition. And FTK imager is more accurate than the Pro discover, and the image verification after the creation of the image is also not checked by the Pro discover tool. The recovery of the deleted file is possible with the help of both tools.

9. Conclusion:

This paper presents the implementation of data acquisition and memory acquisition methods and found that the time taken by the tool to analyze the result is different among the tools. While making the data acquisition of a CD drive, it requires a Fernico device to install to perform the operation in FTK imager. As the paper presents the data acquisition with the help of FTK imager, the author uses a pen drive to acquire the data, so it takes time of 33 minutes to acquire the 32 GB of data and 11 minutes for verifying the image result. But FTK does not have a progress bar to estimate the remaining time; instead, it shows 00:00:00 or nothing. And it also doesn't have the timeline view while data acquisition is performed. We also make the memory acquisition of laptop 4GB RAM, which takes 10 minutes of time for the whole memory acquisition process. This paper also presents the data acquisition of a 32 GB Pen drive with the help of ProDiscover, and it takes 37 minutes to acquire the data, which is 4 minutes more from the FTK imager tool. So we can say that the FTK imager is slightly fast than the ProDiscover tool for performing the data acquisition. ProDiscover has the progress bar to show the remaining time. For the memory acquisitions of Pro discover, it takes 10 minutes for the whole memory acquisition process. So the next goal is to work on more digital forensic tools for data acquisition and memory acquisition and compare their results. This paper will help future researchers to understand which tool is better in the acquisition process.

References:

[1] Muhammad nur faiz and Wahyu adi prabowo. "Comparision of acquisition software for digital forensic purpose." KINETIK, purwokerto, Indonesia 2019.

[2] R. J. Mcdown, C. Varol, L. Carvajal, and L. Chen, "In-depth analysis of computer memory acquisition software for forensic purposes," Journal of Forensic Sciences, Vol. 61, No. January, Pp. 110–116, 2016.

[3] Stefan vomel and Felix C. frieling. "A survey of main memory acquisition and analysis techniques for the windows operating system." Elsevier, Department of Computer Science, Friedrich-Alexander University of Erlangen-Nuremberg, Am Wolfsmantel 46, 91058 Erlangen-Tennenlohe, Germany 2011.

[4]Kambiz ghazinour, Deep m vakharia & Rohit satyakumar. "A study on digital forensic tools." IEEE (ICPCSI), USA 2017.

[5] Ioannis Lazaridis, Theodoros Arampatzis, Sotirios Pouros. "Evaluation of digital forensics tools on data recovery and analysis." Proceedings of the Third International Conference on Computer Science, Computer Engineering, and Social Media (CSCESM2016), Thessaloniki, Greece, 2016.

[6] Varsha Karbhari Sanap, Vanita Mane "Comparative study and simulation of digital forensic tools." International Conference on Advances in Science and Technology 2015 (ICAST 2015)

[7] Brian Carrier. "open source digital forensic tools." London, 2002.

[8] Josiah Dykstra, Alan T. Sherman. "Digital investigation." Elsevier, Cyber Defense Lab, Department of CSEE, University of Maryland, Baltimore County (UMBC), 1000 Hilltop Circle, Baltimore, USA 2012.

[9] Mayank Lovanshi and Pratosh Bansal. "Comparitive study of digital forensics tools." Springer Indore, MP, India 2019.

[10]Raz Ben Yehuda, Erez Shlingbaum, Yuval Gershfeld, Shaked Tayouri, Nezer Jacob Zaidenberg. "Hypervisor memory acquisition for ARM." Elsevier, Finland 2021.

[11] Tomothy vidas. "The acquisition and analysis of random access memory." JDFP, Naval postgraduate school, Monterey CA 2006.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:13:20 UTC from IEEE Xplore. Restrictions apply.

[12] Brain D. Carrier and Joe Grand. “A hardware based memory acquisition procedure for digital investigations.” Elsevier, Department of Computer Sciences, CERIAS, Purdue University, USA 2003.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:13:20 UTC from IEEE Xplore. Restrictions apply.

Conceptual Design of an Unmanned Aerial Vehicle for Mars Exploration

Ishan Mishra, Aayush Kumar, and Vanshaj Malhotra

***Abstract* — Significant technology advances have enabled planetary exploration aircraft to be considered as a viable science platform. These systems fill in a unique planetary science measurement gap, that of the regional scale, near-surface observation while providing a new perspective for planetary discovery. Exploration of Mars using UAV (Unmanned Aerial Vehicle) has been planned for over 25 years by leading space organizations such as NASA. Recent efforts have been able to produce some mature mission and flight system concepts, ready for flight project implementation. There are however si0gnificant numbers of challenges associated with getting an airplane to fly through the thin, carbon dioxide-rich Martian atmosphere. Traditional aircraft design expertise does not always apply to this sort of vehicle, and geometric, aerodynamic, and mission restrictions result in a restricted viable design space. This paper presents the conceptual approach that was taken to design a UAV capable of performing a VTOL (Vertical Take-Off and Landing) in the atmosphere of Mars. The UAV was designed to participate in the International Planetary Aerial Systems Challenge 2021. The UAV could carry a science payload of up to 5 kg (weight on mars).**

***Key words* — Mars Airplane; Science Payload; UAV; VTOL.**

***Abbreviations Used*:**
UAV – Unmanned Aerial Vehicle
VTOL – Vertical Take-Off and Landing
EP – Extreme Pressure
PFPE – Perfluoropolyether
PTFE – Polytetrafluoroethylene
FEA – Finite Element Analysis
NACA – National Advisory Committee for Aeronautics
TCS – Temperature Control System
RHU – Radioisotope Heating Units
CCD – Charge Coupled Device

I. Introduction

For at least 50 years, people have imagined winged spacecraft flying through the Martian atmosphere. The Mars Project, written by Wernher von Braun and originally published in the early 1950s. For almost 25 years, considerable attention has been given to the use of an unmanned aerial vehicle (UAV) as a research platform for exploring Mars. From its unique vantage point, a few kilometres above the Martian surface, an autonomous airplane could return unique science measurements over regional-scale distances. By traversing regional-scale distances at near-surface altitude, planetary airplane observations complement and extend orbital and landed measurements while providing a fresh perspective for scientific discovery.

This paper presents the design of an Unmanned Aerial Vehicle capable of VTOL working under the Martian atmosphere. The UAV is a solar-powered wing body design having the main wingspan of 3.6m and a root chord of 0.5 m. Based on the DOS SAMARA [1] concept the main wing is divided into 2 sections i.e., the inner and outer sections. The outer section is 1 m long and rotates to generate sufficient thrust for VTOL operations. With a full payload of about 12 Kgs, the UAV is capable of traveling 10 Kms having a velocity of 16.66 m/s during the cruise while flying at a max altitude of 7 km from the ground. A propeller located at the back of the plane provides sufficient thrust to counteract the drag forces, putting the plane in steady and levelled flight conditions.

II. Requirements and Considerations

The first and foremost consideration is the nature of the Martian atmosphere. Atmospheric density near the surface of Mars is roughly equivalent to an altitude of 100,000 feet on Earth. It is difficult to create enough lift to maintain the airplane's weight in such a thin atmosphere, thus wing loading must be kept to a minimum.

Owing to the reduced gravity on Mars the weight of the airplane is roughly 38% of its weight on earth. Nonetheless, to attain a reduced wing loading, a considerable focus must be made on reducing airframe and system weight.

Because of the low air density, the cruise velocity is increased, but the flight Reynolds number is reduced. Because of the lower speed of sound on Mars than on Earth, transonic aerodynamic effects are encountered at lower flight speeds. To reduce significant lift losses and/or drag increases associated with low Reynolds number, high Mach number flying circumstances, special consideration was necessary for the aerodynamic design.

Thrust generation is further complicated by low atmospheric density. The lower the density, the less mass available for momentum transfer and the less thrust generated for a given size propulsion system.

III. Mission Layout

Extensive trade studies and design cycles were performed to develop a low-risk mission architecture. The resultant

Submitted on July 11, 2021.
Published on August 15, 2021.
Ishan Mishra
(e-mail: ishanmishra9@gmail.com)
Aayush Kumar
(e-mail: ayushkumar7948@gmail.com)
Vanshaj Malhotra
(e-mail: vanshaj24@@gmail.com)

architecture enabled the satisfaction of the mission requirements with the use of a basic, low-risk flying system design.

The UAV will fly a pre-planned 5 km route from a vantage point 7-8 km above the ground and then return to the base. The UAV could carry a science payload weighing roughly 5 Kgs and an additional payload of 2 kgs that could be connected at the base of the UAV. The science payload would analyse both the Martian atmosphere and surface during flight.

The detailed flight plan is shown in the image.

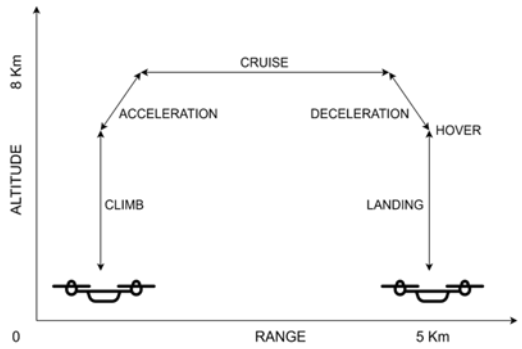


Fig. 1. Mission Architecture.

IV. Modes Of Operation

The hybrid UAV consists of three modes of operation concerning flight stages. VTOL mode for vertical take-off and landing along with the inclusion of loitering as well at a specific altitude and location. The fixed-wing aircraft mode for normal flight operations (horizontal translation).

The transition and tuning mode for switching between VTOL and Fixed Wing mode.

A. VTOL Mode

This phase has been intended for vertical take-off, landing, and hovering operations. During this phase, the outboard wing panels spin generating sufficient thrust to lift the UAV. To save power consumption, the UAV platform automatically switches to aircraft mode after reaching the appropriate height.

B. Transition Mode

The thruster-mounted shaft of the UAV starts and provides a forward thrust. The two outboard wing panels are halted synchronously which causes a halt to the vertical motion of the UAV and the outboard panels now act as a fixed-wing. With an increase in horizontal speed, the fixed wings start to develop lift.

C. Aircraft Mode

During this phase, the UAV has reached a sufficient translational speed and is producing a sufficient lift force to offset its weight. Under these conditions, the platform performs like a typical airplane, with control surfaces such as ruddervators, ailerons, and flaps providing the necessary yaw, pitch, and roll motions. This phase is designed for energy-efficient operation throughout a large chunk of the mission.

V. Design Process

A systematic design approach has been used to design our UAV. The constraints have been analysed first followed by the selection of aerofoils for both wing and tail. A preliminary design of the fuselage, wing, and tail has been developed based on calculations and analysis which has been refined to minimize weight and optimize storage. The final design has undergone several analyses to ensure whether or not the design is adept enough to fly in the challenging Martian atmosphere. Structural and Thermal analyses have been carried out to prevent any kind of mechanical failure during operations. As shown in Fig 2.

VI. UAV Mechanics

All of the materials used in the parts were chosen after careful consideration of different criteria such as weight, hardness, yield strength, tensile strength, and so on.

Owing to the property of gold to block radiation, the parts of the UAV have been gold plated to protect the UAV from extreme levels of UV Radiations.

After extensive investigation and testing of different alternatives, gear lubrication has been assured by a mixture of 50% 802 EP grease and 50% 815Z oil to ensure efficiency even at extremely low temperatures since these lubricants do not dry up quickly. Penzane-3 Pb Np oil and 601 EF have been produced from PFPE oil thickened with PTFE powder [2] for lubricating other components such as bearings, motors, and so on.

The aerodynamic and structural analysis of the UAV has been done considering several factors and the calculations have been made considering the weight of the UAV to be 10 Kgs on Mars.

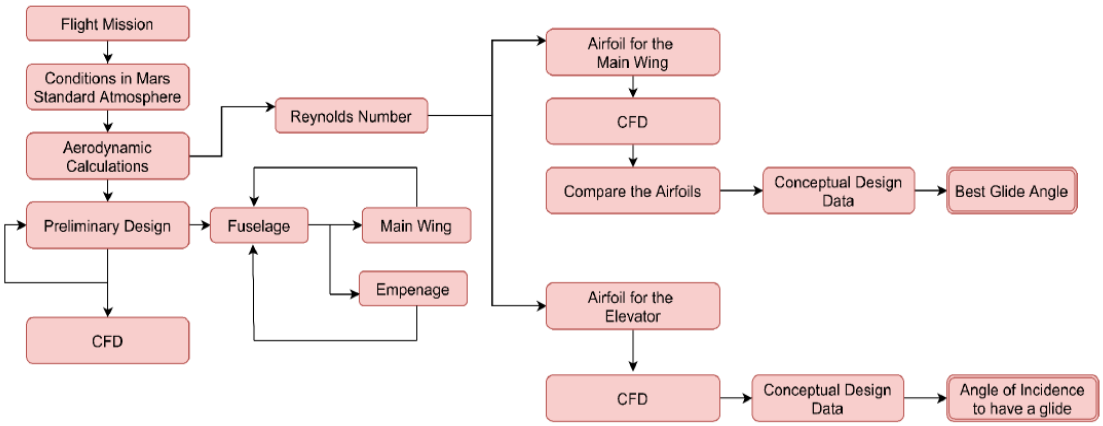


Fig. 2. Design Process.

VII. Geometry

Our plan has mixed the drifting capacity of a helicopter and rapidly expanded the perseverance flying of a fixed-wing airplane. This has helped our UAV to gain the benefits of both into one.

We have used outboard wing panels that are 1 m long and have a root and a tip chord of 0.3 m, that spin to generate thrust to lift the UAV in vertical flight. The wings rotate at 3000 RPM generating a maximum lift of 110 N making them capable of lifting the UAV from the ground. Both rotors rotate in opposite directions hence generating a contour torque preventing the UAV from spinning. During horizontal translation the outboard wing panels lock hence, acting as a fixed-wing. A pusher propeller is located aft of the UAV to provide thrust during the cruise. For deciding the wing geometry, possible planform shapes have been explored extensively, several trade studies between wing area and aspect ratio have been conducted, and the theoretical position of the wing aerodynamic centre has also been restricted to avoid excessive deviations from the needed aircraft centre of gravity. Based on the outcome of these studies, A baseline planform with a 17.4 aspect ratio and a wing size of 1.8 m^2 has been chosen.

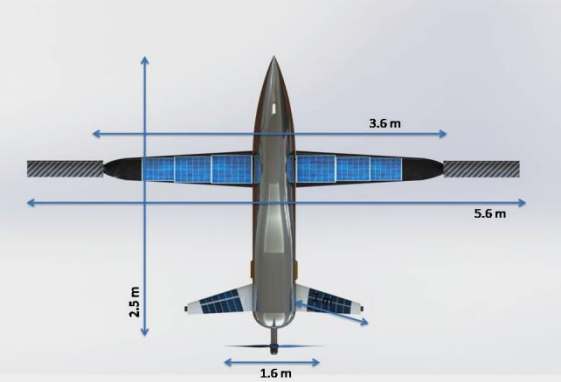


Fig. 3. Dimensions.

On the ground, the plane is supported using a landing mechanism fabricated using Aluminium – 7079 at the front and on the Y-Tail at the aft position. The landing mechanism is connected to the plane using Heim joints providing flexibility to the structure making it capable of landing comfortably on uneven Martian terrain.

VIII. Fuselage

The Fuselage is an integral part of the UAV. It has been designed to maximize storage capacity while minimizing weight. The fuselage houses the Flight Electronics, Science Module, and everything that has to be incorporated in or on the UAV. The Fuselage is 1.5 m long and about 0.6 m wide making it adept to cache all the components appropriately, enabling a compact and efficient design. The Fuselage has been fabricated using Mylar and AS Graphite-Epoxy composite [3]. This material provides an excellent strength-to-weight ratio and could easily withstand both temperature extremes on Mars. Ease of Serviceability has been kept in mind while designing all the components making all subsystems accessible to the astronaut during maintenance. Openings have been provided both under the fuselage as well as on the rear end which makes it easier to access the electronics without dismantling the wings.



Fig. 4. Lateral View.

IX. Aerofoil Selection

Several Airfoils including Eppler 203, NACA 66(3)-418, RG-15[4], i.e., those capable of operating under Low Reynolds Number and High Mach Number (0.6-0.8) on the Martian atmosphere have been considered and after performing extensive analysis of those airfoils considering various key factors such as density, pressure, gravity, etc.

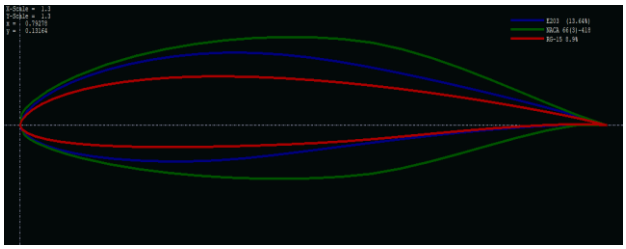


Fig. 5. Aerofoil Comparison.

RG-15 has been selected for our design. The aerofoil provides a minimum drag coefficient as well as a maximum lift coefficient, which means the highest lift-to-drag ratio; also, it provides a low pitching moment coefficient. These characteristics empower our UAV wing with noteworthy thrust and negligible drag, subsequently diminishing the power consumption and force utilization for an efficient flight. The aerofoil which has been selected for our UAV guarantees a positive Coefficient of Pressure enabling us to achieve a Lift Coefficient as high as possible. An angle of attack of 8 deg. has increased the aerodynamic force and has streamlined power, along with giving a positive moment about the Aerodynamic Centre.

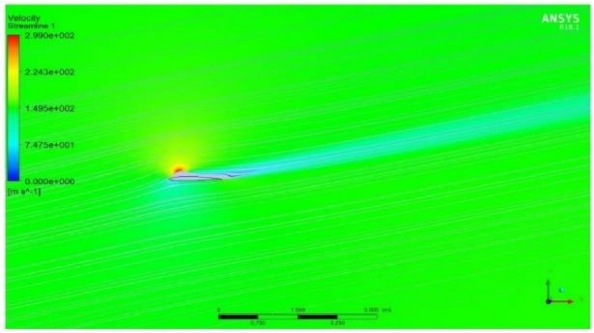


Fig. 6. Pressure and Velocity Graph of RG-15.

X. Rib Design

Morphed wing configuration has been for the UAVs fixed wings, which consists of three main components: aluminium profiles, low-temperature foam, and silicone elastomers.

A. *Low-Temperature Foam*

It has been used to fill the gap between the aluminum profiles. A protective layer of low-temperature silicone elastomers has been applied to aluminum profiles and foam [5].

B. *Skin Material*

Silica Elastomer has been used as the material for the skin hence making the morphed wing adaptable and flexible so that the material is handily twisted, while as yet having the solidarity to convey the streamlined loads of the UAV. Likewise, the material has been supplied with an abrasion-resistant coating so that the material is not harmed when liable to evolving conditions. After the material has been converted to an alternate wing shape, the material recuperates its original size. The "skin" of the wing additionally upgrades the construction's exhibition. It has been produced using covering portions of adaptable material, layered to some degree like quills or fish scales, taking into account the pieces to get across one another as the wing flexes, while as yet giving a smooth external surface.

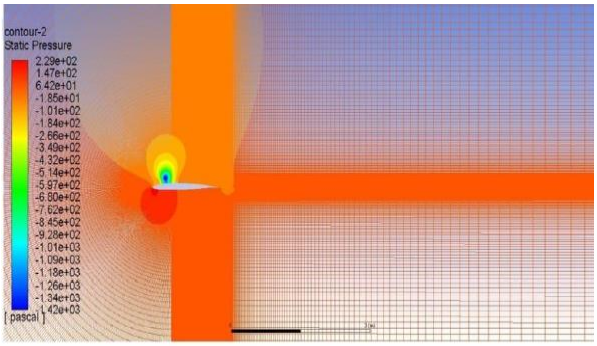


Fig. 7. Rib Structure.

C. *UAV Morphing Plan [6]*

The plan will affect the flying performance of our UAV by the dramatic change of wing shapes. Such plans take out the requirement for various, costly, mission-explicit aerofoils. Morphing aerofoils can change their shape for various flying performances. Such changes can be either incremental change to facilitate flight control. The key to the successful morphing of our wing was geometric scaling to match a particular performance requirement. The morphing would be carried out using 4 servo motors placed within the fixed-wing and connected to the spars. These ensure precise morphing to facilitate the required aerodynamic characteristics.

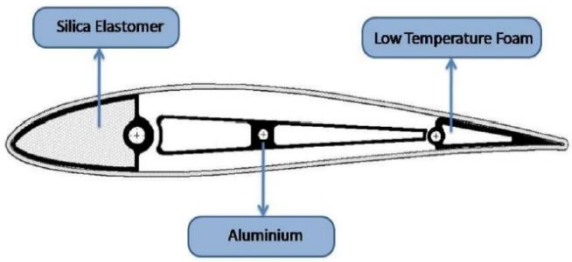


Fig. 8. Skin.

XI. Empennage Design

The Empennage provides for trim, stability, and control and is designed to operate normally at only a fraction of lift potential. Y- Tail configuration has been chosen for reducing drag interference, avoiding the complexities of ruddervators, and keeping the propeller off the ground when compared to conventional tail. The horizontal tail provides better stability and the vertical tail reduces aerodynamic sideslip and provides directional stability. NACA 0009, has been chosen for our UAVs tail which guarantees pitch control and longitudinal stability since it is an asymmetrical airfoil. This airfoil also provides a negative lift with an increased angle of attack, resulting in a negative moment, allowing the UAV to remain naturally stable.

TABLE I: Empennage Design Parameters

Parameters	Horizontal Tail	Vertical Tail
Root Chord	0.4188 m	0.25m
Tip Chord	0.1637m	0.15m
Total Surface Area	0.74 m2	0.344m2
Wing Span Length	2.5408m	0.3m
Wing Aspect Ratio	8.71	1.60

XII. Propellor Design

The propeller blade has been designed for the lowest weight possible. Mylar, a light plastic material, has been chosen as the primary component of the skin structure. Because of its excellent strength-to-weight ratio, graphite-epoxy composite [7] was chosen as the material for the skin-supporting structure. The baseline design consists of a single hollow spar with rib elements to support the Mylar skin. To avoid aeroelastic instabilities, the blade's centre of gravity has been restricted to be coincident with the quarter chord axis position. NACA 16-012 aerofoil has been selected for our propeller. With a 70° twist angle and desired minimum thickness of 12%, our propeller ensures a high turning rate with a motor of 3000 RPM. The blade design is inherently smoother with reduced tip speed and minimal noise. This design marks constant thrust along the entire length of blades thereby maintaining the pitch to diameter ratio of 1:1 and making it more efficient.

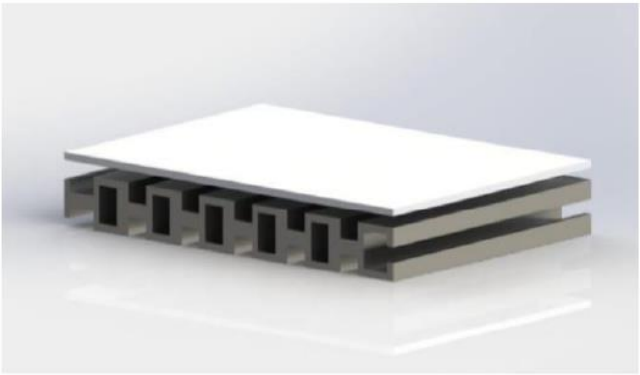


Fig. 9. Weight Distribution.

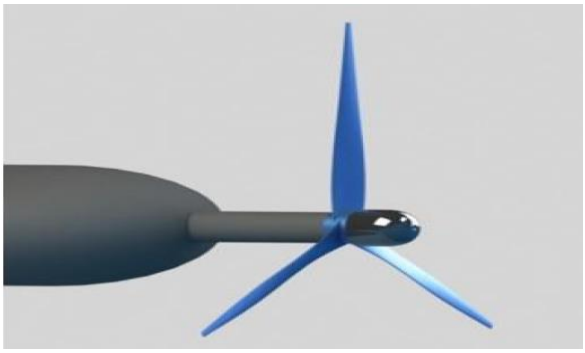


Fig. 10. Propellor.

XIII. Weight Distribution and Balancing

The well-known Boeing-Vertol equations were used in preliminary investigations to determine the weight breakdown. This algorithm takes size, gross weight, and other basic parameters and compares them to historical patterns. The output of these equations includes more detailed weight information on UAV subsystems.

TABLE II: Weight Division

S.no	Sub-System	Weight (Kg)
1.	Structural	4.0
2.	Batteries and Propulsion	1.2
3.	Electrical	0.3
4.	Science Payload	4.7
	TOTAL	**10.2**

The design has been optimized using the data obtained to have a maximum scientific payload capacity without exceeding our weight. Our final weight estimates slightly deviated from the traditional Boeing-Vertol formulae due to the following reasons:

- Our fuselage is lighter than anticipated since our plane is unmanned. Composite materials have also been used throughout instead of traditional heavier metals.
- Our wings and blades have been made using composite materials hence are lighter than traditional designs. Hence, they were lighter than predicted.

XIV. Environmental Controls

A. *Temperature Management*

To make our UAV adaptable to the fluctuations in Martian temperatures from 20 °C to a freezing -15 0°C we have introduced a wire meshing network in the main UAV body. The meshing is the primary heating system for the Electronics and Science subsystems. The network is connected to the centralized power system via a relay. The meshing is made of electrically insulating and thermally conductive diamond coating on copper coils [8] thus providing electrical safety to the components without compromising on thermal parameters of the devices used.



Fig. 11. Heat Switch.

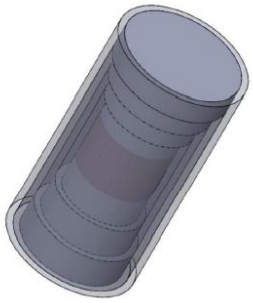


Fig. 12. RHU.

In addition to the meshing network, a robust Temperature Control System (TCS) has been incorporated. TCS is inclined towards sustainable temperature management without compromising any power consumption. It consists of 8 Radioisotope Heating Units (RHUs) [9] spread across the plane. RHUs have an excellent capability to produce heat and additionally provide the benefit of reduced electromagnetic interference to their electrical heating counterparts. It provides a constant supply of 1W heat without adding on a burden on the power cache as they provide heat through radioactive decay. The placements of RHUs have been criticized due to the difference in the amount of heat required by individual components.

Adequate temperature range for each component is maintained using heat switches and insulation as discussed below:

Maximum heat insulation has been provided for the UAV to ensure it withstands cold Martian. All electronic components are placed within this insulated box known as Heating Module. It has been further thermally regulated using graphene tubing [10] perforated through the box which conducts heat through it to the base of the payload carrying mechanism. It prevents our UAV from overheating during operations on warm Martian days.

The radiators have been designed such that, in the worst scenario of temperature variation they maintain the optimal working temperature in the Heating Module. To prevent heat loss during the night, the conduction tubes would be connected to the radiator plates through Shape Memory Heat Switches [11]. These are metallic materials that can return to a previous shape or size when subjected to the right thermal conditions. These switches can operate within a temperature range of roughly -200 C to +100 ºC, making them suitable for the Martian atmosphere.

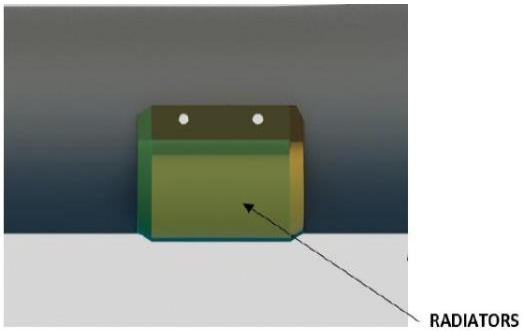
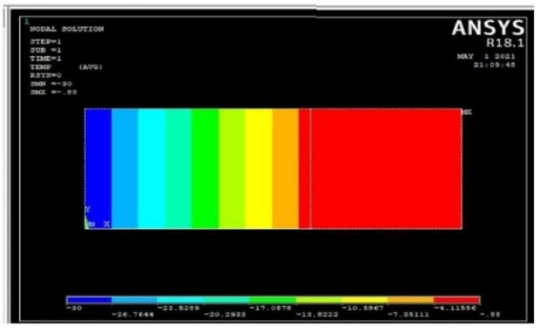


Fig. 13. Radiators.

To trap heat inside the UAV the inside of the UAV would be flushed with CO2 gas. Also, a layer of solid silica aerogel [12] would be placed around the electronic box and the motors. Both carbon dioxide and silica aerogel are excellent insulators of heat. These prevent any unwanted heat loss from the plane's surface. Graphene has been employed as the material for making the tubes as it has an excellent conduction and convection potential, hence making it an ideal material for making the tubes. To allow efficient heat transfer between radiators and the Martian atmosphere the radiators are made of copper which is an excellent conductor of heat.



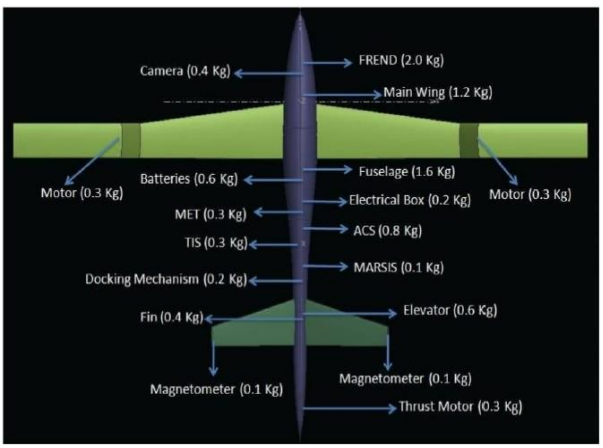


Fig. 14. Thermal Analysis of Fuselage.

B. *Dust Protection*

Mars is infamous for intense dust storms, which sometimes kick up enough dust to be seen by telescopes on Earth. Individual dust particles on Mars are very small and slightly electrostatic, so they stick to the surfaces they contact, and consequently, the possibility of dust settling on and in machinery is a challenge for the mission. To prevent this from happening the entire exposed surface of the UAV is covered with a dust repellent layer of CVD Nano-Coating [13] that prevents the UAV surface from wear due to the bombardment of fast-moving dust particles. To protect the cameras and lenses from damage due to dust and wind, they are equipped with lens-protective covers.

XV. Science Payload

The science payload for the UAV comprises the following components:

A. *Planetary Surface Analyzer (PSA)*

It is a CCD Camera and electronics kit with a dual optic path and a single focal plane scope. It would be used for providing regular summary and low-goal imaging of Mars to examine the climatic changes.

B. *Calorific Imaging Spectrometer (CIS)*

It is a plane reflection spectrometer that would be used for the analysis of the Martian atmosphere and to measure the surface composition of CO_2, N_2, Ar, O_2, etc.

C. *Atmospheric Data System (ADS)*

It will provide a comprehensive measurement of the Martian near-surface atmosphere. Also, it would be used to measure the wind speed and direction.

D. *Infra-Red Spectral Chemistry Suite (ISCS)*

The primary goal of the ISCS instrument would be to detect minor atmospheric gases along with sensitive detection of minor gases and Isotopes to address global scientific issues related to Mars.

E. *High Precision Neutron Detector (HPND)*

It would help gather more information about the geology of Martian soil. It would also be used to create high-resolution maps of hydrogen-abundant regions on Mars.

F. *Subsurface Sounding Radar and Altimeter (SSRA)*

The primary goal of SSRA would be to map the distribution of water and ice in the upper layers of Martian Crust. It also acts as an altimeter.

G. *Magnetometer*

It will study Mars' magnetic and electric fields as well as its plasma environment.

XVI. Conclusion

We presented the conceptual aerodynamic design of a UAV in this article capable of performing a VTOL (Vertical Take-Off and Landing) in the atmosphere of Mars. The UAV is capable of traveling over a range of 10 km per flight at an altitude of 7-8 km above the surface. It is equipped with Cameras and Sensors that would be used to analyze the surface and atmosphere of Mars providing us with a better understanding of the Planet. It is capable of withstanding harsh Martian climate owing to its Temperature control and dust protection systems. Such UAVs will have a crucial role in future Mars missions as they provide us with a better overview of the Planet.

XVII. Future Scope

A. Surface Mapping

The UAV could be used for mapping the surface of Mars. The data so obtained could be used for planning further rover missions or to find a suitable site for humans to settle.

B. Transportation

The UAV could be used to transport essential items over a larger range faster.

C. Exploration

The UAV could be used to Explore hard-to-reach places and provide the required data for further analysis.

D. Meteorology

The UAV could be used to monitor regular changes in climatic conditions. Providing a better understanding of the Martian atmosphere.

References

[1] William J. Fredericks, "Conceptual design of a VTOL Unmanned Aerial Vehicle with 24Hr Endurance".
[2] Robert L. Fusaro; "Lubrication of Space Systems".
[3] Wolf Elber; "Toughening of Graphite-Epoxy Composites by Interlaminar Perforated Mylar Films"; NASA March 1978.
[4] California State University Northridge; "Project Ares II: High-Altitude Battery Powered Aircraft"; Proceedings of the NASA/USRA Advanced Design Program 7th Summer Conference.
[5] Oliver Schorsch, Andreas Luhring, Christof Nagel et al., "Polymer Based Morphing Skin for Adaptive Wings".
[6] Hangil You, Seongik Kim, et al.; "New Concept for Aircraft Morphing Wing Skin: Design, Modeling, and Analysis". DOI: 10.2514/1.J058102
[7] Rudy Lukez; "The use of Graphite/Epoxy Composite Structures in Space Applications".
[8] ", Daichi Nakajimaa, Hiroki Kuwabara; "Diamond-like carbon coating for effective electrical insulation of Cu and Al wiresNASA Article on "Radioisotope Heater Unit".
[9] Yifeng Fu; "Graphene related materials for thermal management" et al 2020 2D Mater. 7/01/2001.
[10] K. Lankford; "Heat Switches" Starsys Research Corporation, Boulder Colorado.
[11] J.E. Fesmire; "Aerogel Insulation Systems for Space Launch Applications."
[12] A. Delfini, A. Vricella, R. Bueno Morles." CVD nano-coating of carbon composite for Space materials." Procedia Structural Integrity Volume 3, 2017.



Ishan Mishra was born on 14th February 2001 in Dharan, Nepal. He is currently pursuing his B. Tech Degree in Mechanical Engineering from Delhi Technological University and is expected to complete his B. Tech in 2023.
He is currently working as the Head of Design in Team Inferno DTU which is a student-based technical team working on the design and fabrication of Mars Rover Prototypes. His areas of research include Robotics, Design of Unmanned Aerial Vehicles, Design of Mars Rovers, and Mars Explorations. He has presented a paper on "Smart Manufacturing: Opportunities and Challenges overcome by Industry 4.0" in the International Conference on Advanced Manufacturing and Materials Processing (CAMMP 2021) organized by MNIT Jaipur and a Journal paper on "Conceptual Design of a 7 DoF Robotic manipulator for Martian Applications" in IJISET Vol 8 Issue 6, 2021.



Aayush Kumar was born on 14th August 2001 in Gurugram, Haryana. He is now studying his B.Tech Degree in Mechanical Engineering with minors in robotics and automation at Delhi Technological University, and he is anticipated to complete his B.Tech in 2023.
He is presently the Mechanical Lead of Team Inferno DTU, a student-led technical team tasked with the design and fabrication of Mars Rover prototypes. His scientific interests include robotics, the design of the Mars Rover prototype, a 6 Degree of Freedom robotic arm manipulator, and his ongoing project include Conceptual design of fixed-wing UAVs unmanned aerial vehicles for both terrestrial and extra-terrestrial explorations his expertise lies in the field of fluid dynamics and finite element analysis.



Vanshaj Malhotra was born on 24th May 2001, in Delhi India. He is currently pursuing his B.Tech Degree in Mechanical Engineering with Minors in Robotics and Automation from Delhi Technological University and is expected to complete his B.Tech in 2023.
He is currently working as a Team Leader in Team Inferno DTU which is a student-based technical team working on the design and fabrication of Mars Rover Prototypes. His areas of research include Robotics, Design of Unmanned Aerial Vehicles, Design of Mars Rovers, and Mars Explorations. He has published a conference paper on "Role of Robotics in Mass Manufacturing and its Associated Advantages and Disadvantages" CAMMP 2021 organized by MNIT Jaipur and a Journal paper on "Conceptual Design of a 7 DoF Robotic manipulator for Martian Applications" in IJISET Vol 8 Issue 6, 2021.

2021 Fourth International Conference on Computational Intelligence and Communication Technologies (CCICT) | 978-1-6654-2392-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/CCICT53244.2021.00030

Convolutional Neural Network for Covid-19 detection from X-ray images

Anju Jain
Dept. of *Computer Science & Engineering*
Guru Jambheshwar University of Science and Technology

Hisar,Haryana,India
anju_jain7@rediffmail.com

Saroj Ratnoo
Dept. of *Computer Science & Engineering*
Guru Jambheshwar University of Science and Technology

Hisar,Haryana,India
ratnoo.saroj@gmail.com

Dinesh Kumar
Dept. of *Electronics Computing & Engineering*
Delhi Technological University

Daulatpur,Delhi,India
dinesh_chutani@yahoo.com

***Abstract*- Covid-19 pandemic has crumbled the health systems of the nation's world over. In such a scenario, quick and accurate detection of coronavirus infection plays an important role in timely referral of physicians and control transmission of the disease. RT-PCR is the most widely used method for identification of coronavirus disease 19 patients, but it is time consuming and takes two to three days to deliver the report. Researchers around the world are looking for alternative machine learning techniques including deep learning to assist the medical experts for early Covid-19 disease diagnosis from medical pictures such as X-ray films and CT scans. Since the facility for chest X-rays is available even in smaller towns and is relatively less expensive, it would be useful to design machine learning methods for proving initial Covid-19 detection from chest X-rays to contain this pandemic. Thus, in this work, we propose a Convolutional Neural Network (CNN or ConvNet) for the finding of presence and absence of Covid-19 disease. We compare the CNN model with traditional and transfer learning-based machine learning algorithms. The proposed CNN is accurate compared to the traditional machine learning algorithms (KNN, SVM, DT etc.). The suggested CNN model is almost as accurate as the classifiers based on transfer learning (such as InceptionV3, VGG16 and ResNet50) despite being simple in terms of number of parameters learnt. The CNN model takes less training time.**

Keywords— Covid-19 disease diagnosis, Transfer Learning, Deep Learning, Convolutional Neural Network, Chest X-ray image, Pneumonia

I. INTRODUCTION

A new coronavirus, initially originated from Wuhan city in China, turned into a pandemic almost around the globe. Due to the unavailability of vaccines against the new disease Covid-19, early diagnosis was of extreme importance. This was needed to immediately isolate the suspected persons for preventing the further spread of infection. The current only method of medical checkup for Covid-19 is reverse transcription polymerase chain reaction (RT-PCR) [1]. However, the accuracy of RT-PCR on swabs deep in the nasal cavity, the front of the nose or the throat has been reported to be only 30% to 60%, which is very dangerous [2]. The commonly used tool for pneumonia diagnosis is X-ray photograph or computer-assisted tomography (CT) scans. Both tools have high sensitivity for diagnosing Covid-19 induced pneumonia [3]. X-rays show visual indices that correlate with Covid-19 and capture inflammation in the lungs. Medical practitioners initially prefer Chest X-ray pictures more than CT-scan for pneumonia diagnosis since it is easily available, cheaper, faster and exposes the patient to less radiation[4] [5] [6]. Because there is an enormous shortage of trained radiologists, computerized methods can increase the speed of timely detection of Covid-19 disease.

Machine learning provides efficient tools for solving disease diagnosis problems. Predicting Covid-19 pneumonia severity from chest film using expert system machine intelligence may be an option. Traditionally a two-step procedure is followed for medical image classification, namely features extraction and classification. The end-to-end deep learning structure can predict Covid-19 directly from the original image without extracting features. CNN-based models based on deep learning are superior to traditional artificial intelligence techniques in several areas such as computer vision, facial recognition and medical image analysis. CNN-based models have been used for a varied problem such as classification, regression, segmentation and image optimization [7].

In the situation of the Covid-19 pandemic, we seek to improve the prognostic forecasts to manage the patient care. In this research, we attempt to apply Convolutional neural networks (CNN's) for diagnosing Covid-19 by X-ray films. We compare the performance of CNN with some popular pre-trained ImageNet models for Covid-19 diagnosis from X-rays. The pre-trained models, which come under the domain of transfer learning, are considered accurate. However, the pre-trained models are complex as compared to simple CNN models. The contribution of this research is to compare the performance of ConvNet with pre-trained models in the context of diagnosing coronavirus disease 19 from chest X-ray films.

The organization of the rest of this article is as follows: Section 2 describes essentials of CNN and transfer learning. Section 3 provides literature review of CNN and transfer learning-based machine learning algorithms for Covid-19 disease diagnosis. Next section presents the materials and methods used for this study. Section 5 provides the experimental studies and observations, and finally the paper is concluded in the last section.

II. BACKGROUND DETAILS

Since CNN has shown promising results for classification tasks, we have used CNN for Covid-19 disease diagnosis and compared its performance with state-of-the art deep transfer learning models in this paper. Description of CNN and transfer learning is as follows:

A) CNN or ConvNet: Convolutional neural networks are a specialized type of multi-layer neural network mainly used in image processing and computer vision domain for feature enrichment. These are designed to detect visual patterns directly from pixel images with minimal preprocessing. They can extract topological properties from an image. In CNN, there is no need to extract the features by using other signal processing techniques. Due to its multilayer structure, CNN has the ability to extract low, medium and high-level features. High-level features are a combination of lower and middle-level features. CNN uses a series of filters to process the raw pixel data of the image to extract and learn advanced features that the model can use for classification. CNN mainly consists of four types of layers: convolutional layers, pooling layers, fully connected layers and output layers. A typical CNN architecture is shown in Figure 1. The convolutional layer is made up of a

978-1-6654-2392-2/21/$31.00 ©2021 IEEE
DOI 10.1109/CCICT53244.2021.00030

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:11:41 UTC from IEEE Xplore. Restrictions apply.

convolutional kernel that works by dividing the image into smaller pieces, usually called receptive fields. This layer uses a matrix filter and performs convolution operations to find patterns in the image. Pooling layer regularly follows the convolution layer. It combines a set of values into a single value, resulting in a reduction in the size of the function map. It converts the representation of the common feature into valuable information, retains useful information, and eliminates irrelevant information. Fully connected layer connects each neuron in one layer to each neuron in another layer. In principle, it is the same as the traditional multi-layer perceptron neural network (MLP). It takes the output of the previous layers, "flattened" it, and then converts it into a single vector, which can be used as the input for the next step. The output layer might produce likelihoods of an image belonging to a particular class. This layer uses the sigmoid function as the more generalized nonlinear activation function for binary classification. The first two layers are used for feature maps while the last two layers, which are fully connected, are used for classification.

.

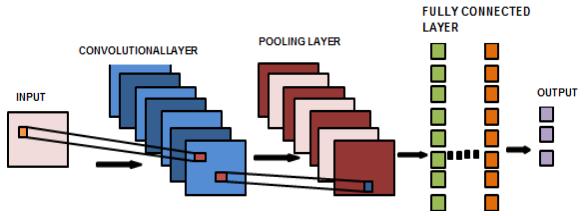


Fig. 1. A Typical CNN Architecture

B) Transfer Learning Approach

In transfer learning, a model learns from one type of problem and then the learning is applied to solve a different but related problem. For example image classification models trained on the ImageNet database with stored weights can detect Covid-19 on the data set used in this paper. Transfer learning is primarily suited to issues where less training data can be used early on to train the model, such as medical image classification. The lower layers of the network have already been trained, the network has learned the general features such as shape and edge of the image etc. and only the higher layers of the network need training. The main advantages of integrating transfer learning with CNN are saving training time and better performance of neural networks even with small amounts of data.

In our study, such pre-constructed models are used as feature extractors. Only the last layer of CNN is tuned according to the number of categories in which the dataset should differ. We have evaluated the performance of three popular pre-trained models on ImageNet database such as InceptionV3, VGG16 and ResNet50.

III. LITERATURE REVIEW

The 2019 new coronavirus (Covid-19) shows many exceptional features as it is very dynamic and continues to evolve [8] [3]. Although the diagnosis can be confirmed by RT-PCR, CT scan and chest X-ray films of patients with pneumonia have also been widely used to detect and examine suspected or confirmed cases of Covid-19. These images show a pattern that is sometimes undetectable for the human eye [9].

Diagnosing Covid-19 on chest radiographs using CNN is one of the most commonly used methods. In [10], authors suggested 121 layers of convolutional neural networks based on chest X-ray dataset and the results were superior in comparison to four practicing academic radiologists. Authors in [11] have proven the useful application of deep training models based on the tailored COVIDX-Net system for Covid-19 X-ray films classification and reported an accuracy of 90%. Wang et al. [12] introduced Covid-Net, a deep ConvNet design to detect coronavirus disease 19 from X-ray pictures and compared the results with VGG 19 and ResNet-50 transfer learning models. At the end of the study, 93.3% normal accuracy was reported. In [13], Covid-ResNet was used as a classifier for Covid-19 disease detection and three other types of infections. The authors reported superior performance compared to Covid-Net with an accuracy of 96.23% and sensitivity 100% for Covid-19 classification. Deep CNN models were used to abstract features from Covid-19 X-ray films [14]. The extracted features from these deep models were classified by SVM. The proposed hybrid model, ResNet 50 and SVM, statistically outperformed other CNN models with accuracy of 95.38%. Abbas et al. [15] proposed a deep CNN called DeTraC for the same. It could detect any irregularity in the image dataset by the class decomposition mechanism. The results of the experiment showed an accuracy of 93% by DeTraC in detecting X-ray films of Covid-19 of healthy and infected cases.

In addition to chest X-rays films, studies were also performed for the detection of coronavirus disease 19 using CT scan. In [16] [17], authors developed a novel weakly supervised deep structured learning model to diagnose Covid-19 disease using upper body CT scans without the need for lesion interpretation for training. The proposed model achieved an average of 88.55% accuracy and 93% sensitivity features respectively, thus providing a quick way to detect Covid-19 patients. However, the main disadvantages of using computed tomography are high radiation, high dose to the patient, and the cost of the examination [18]. In contrast, conventional X-ray machines are capable of producing two-dimensional (2D) projection images of a patient's lungs in all hospitals and clinics. In general, X-ray modulation is the choice of radiologist for the diagnosis of diseases and is used to detect Covid-19 in an average number of patients [9] [19]. Thus, this work focuses only on the usage of X-rays in potential patients with Covid-19.

Another popular CNN architecture learning strategy is to transfer the knowledge learned from a task-driven network to a new task [20] [21]. Pre-trained CNN models such as VGG19 and MobileNetV2 have achieved quite high classification accuracy (97.82) in detecting Covid-19 from X-ray pictures [22]. Similarly, [23] proposed an alternative fast screening method nCOVnet using VGG16 top layers as a bottom model and adding five custom layers as top model showed an overall accuracy of 88.10% in disease diagnosis by X-ray imaging.

Upon review of the related work, it is clear that pre-trained CNN models, although fast and easy to use, achieve higher accuracy at the expense of computational and architectural complexity. Thus, in this study we use CNN architecture which is trained from scratch but has low architectural

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:11:41 UTC from IEEE Xplore. Restrictions apply.

complexity. The CNN model is less complex, fast in speed and easy to deploy. In the study, a total of 539 chest X-ray pictures were examined for binary classification (Covid-19 and non-Covid-19). All the images have been preprocessed to be given as input to CNN.

IV. MATERIALS AND METHODS

A. The Dataset

The dataset used in this study is the coronavirus disease 19 chest X-ray dataset from two online open-source communities –Cohen [24] and Kaggle [25]. Upper body X-rays of patients with Covid-19 were obtained from a GitHub repository shared by Dr. Joseph Cohen. This repository contains 305 chest X-ray pictures of Covid-19 infected people. X-rays of the chest of healthy individuals were obtained from the Kaggle repository. It contains 234 chest radiographs of healthy individuals. All the images were examined by a board of certified doctors. The dataset is split into three parts namely train (70%),test (15%) and validation (15%) dataset.

B. The Proposed Network Architecture

Unlike the complex architecture of pre-trained models, here we have proposed a simple architecture of CNN and trained it from scratch. The proposed method is based on enhanced CNN learning architecture for classification of chest X-ray films into images with or without Covid-19. All the images have been resized to 240 x 240. The images are first preprocessed to remove noise like blur, high contrast etc. Pre-processing involves several operations on images such as gray scale conversion, image segmentation and morphological operations to remove the unnecessary part of X-ray pictures.

The projected CNN model consists of several convolutional layers, which provide efficient feature maps as input to the primary layer. The model has three convolutional layers, then three pooling layers and two fully connected layers, a flattening function and a sigmoid layer as shown in Figure 2. The model uses the RELU activation function as it reduces the likelihood of vanishing gradient sparsity.

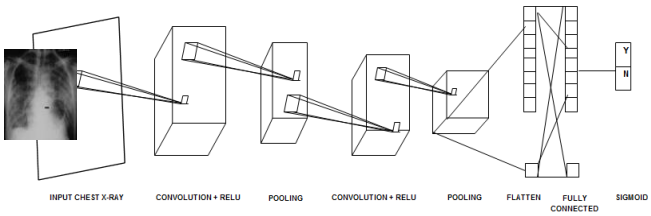


Fig. 2. The Proposed CNN Architecture

C. Performance Metrics

The performance of the deep learning model is evaluated on test data by measuring accuracy, precision, recall and area under ROC.

D. Tools

All the computational experiments are conducted in Python 3 Google Compute Engine Backend (GPU). GPU facilitates faster training of the deep learning and transfer learning models as compared to CPU. Each model is subtly tuned for 24 epochs. The model is trained with a batch size of 32 and ADAM optimizer with step size parameter of 0.0001 which is used to optimize the loss function.

V. RESULTS AND DISCUSSION

Here, we contrast the deep learning model for Covid-19 diagnosis with some of the popular transfer learning models such as InceptionV3, VGG16 and ResNet50. Behavior of models is analyzed with respect to loss/accuracy on training and validation dataset. Figure 3 shows the loss and accuracy plots of all the four models trained for 24 epochs.

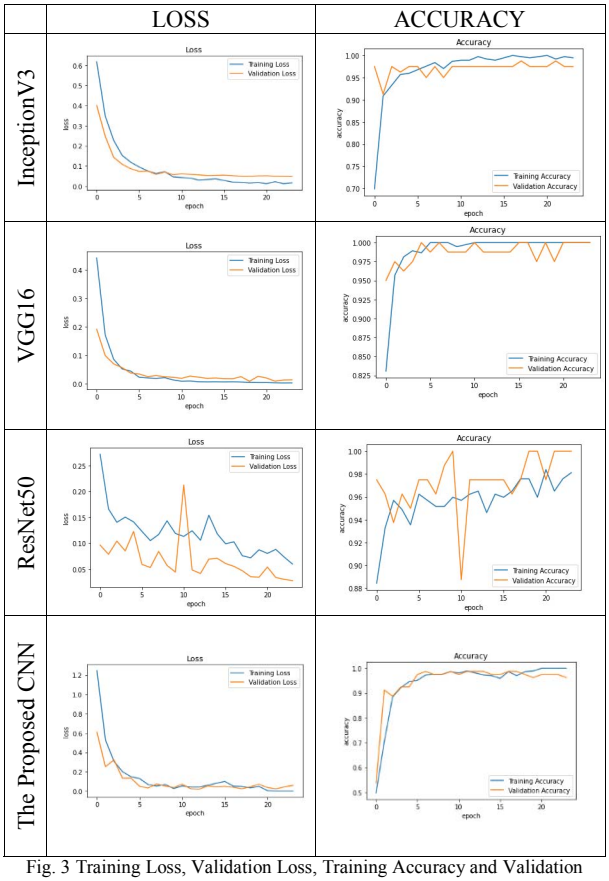


Fig. 3 Training Loss, Validation Loss, Training Accuracy and Validation Accuracy of Learning Models across the epochs.

Figure 3 shows that the ResNet50 model faces the problem of overfitting as train loss continues to decrease with experience and validation loss decreases first and then increases again. Large gap between the two shows that the model does not perform well on the unknown Covid-19 dataset. The proposed CNN model learning curve shows good fit with a small gap between the training loss and validation loss. CNN's performance is comparable to that of the InceptionV3 and VGG16 models. To know the exact number of correctly and incorrectly diagnosed cases, confusion matrix and accuracy of transfer and deep

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:11:41 UTC from IEEE Xplore. Restrictions apply.

structured learning architectures for the Covid-19 disease diagnosis on the test images are shown in Figure 4.

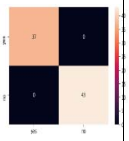
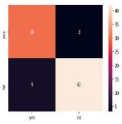
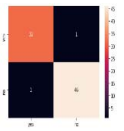
	InceptionV3	VGG16	ResNet50	Proposed CNN
Confusion Matrix				
Accuracy	0.97	1.0	0.94	0.97

Fig. 4 Confusion Matrix and Accuracy of transfer and deep learning architecture for the Coronavirus disease 19 classification

In addition to accuracy, other metrics such as precision, recall and ROC must also be reported for evaluating the performance of a classifier in medical diagnosis. Receiver Operating Characteristics (ROC) curve is used to give overall model performance as the curve is composed of false positive rate and true positive rate. Higher the area under the curve better is the performance of the model on unseen data. The ROC curve for all pre-trained models is shown in Figure 5.

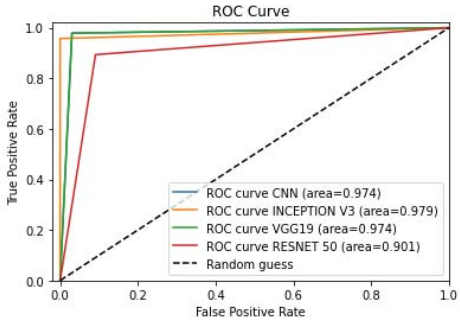


Fig.5 Four CNN ROC Curve for test set

All CNN architectures give higher AUC as compared to the Resnet model. Further, we give a Precision-recall curve to provide the true positive rate as a function of positive predictive rate. Precision-recall curves for all four models are shown in Fig. 6.

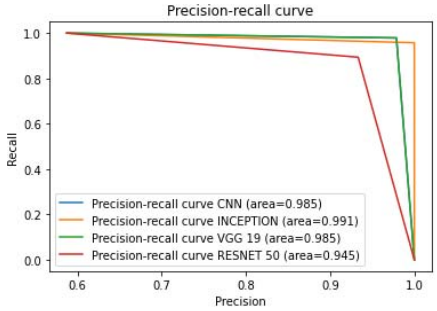


Fig.6 Four CNN Precision-recall Curve for test set

We have also compared the proposed CNN to other traditional machine learning algorithms. Fig. 7 gives the accuracy and standard deviation (in bracket) of several other traditional machine learning algorithms (Linear Regression (LR), Linear Discriminant Algorithm (LDA), K-Nearest Neighbour (KNN), Decision Tree (DT), Naïve Bayesian (NB) and Support Vector Machine (SVM)) on Covid-19 dataset for disease detection. The proposed CNN and the pre-trained models perform better than these machine learning algorithms given in Fig. 7.

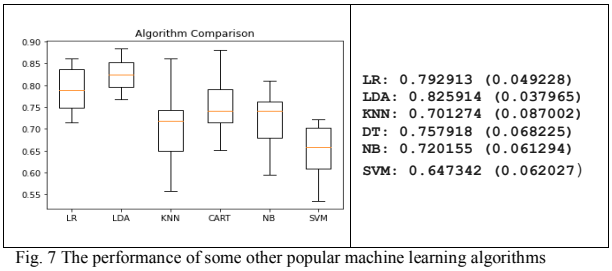


Fig. 7 The performance of some other popular machine learning algorithms

Further, the CNN model is more generalized than transfer learning models (Table 1). It is neither overfit nor underfit. The transfer learning models use excessive resources and complexity. These models use a large number of parameters which results in more training time for the model as shown in Table 1. There is no significant difference between the accuracy of CNN and transfer learning models, but the size of the transfer learning modes is usually big, and this makes these models difficult to deploy. The CNN model is trained from scratch; as a result, it uses a smaller number of parameters and takes significantly less training time except VGG16.

TABLE 1 TOTAL PARAMETERS AND TOTAL TIME TAKEN BY LEARNING MODELS

	InceptionV3	**VGG16**	**Resnet50**	**CNN**
Total Parameters	22,098,337	23,269,057	40,398,337	5,959,617
Elapsed Time	14min 42sec	01min 51sec	9min 54sec	1min 40sec

VI. CONCLUSION

This research proposes a ConvNet to diagnose Covid-19 disease from chest X-ray films. The CNN model has shown a useful and robust response to cases of Covid-19. Multiple convolutional layers of CNN make it effective to extract features of Covid-19 from chest X-ray pictures. We have compared CNN with three popular already trained CNN models. The ConvNet model noticeably outperforms the traditional machine learning algorithms for novel coronavirus disease diagnosis. The model is reliable, its accuracy is competitive, gives outcome quickly and requires reduced computational effort to determine the presence of Covid-19 induced pneumonia. The results show that CNN

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:11:41 UTC from IEEE Xplore. Restrictions apply.

model is a promising option to assist the radiologists to rapidly decide on X-ray films in diagnosing Covid-19. Machine learning algorithms such as CNN can significantly lessen the workload of medical practitioners and increase their efficiency in diagnosing coronavirus disease 19 pneumonia from chest X-rays which are less expensive and are easily available in smaller towns. More experimentation is required to further validate the finding of this research on the performance of CNN and pre-trained models on large repositories of X-ray films and CT scan pictures.

References

[1] W. Wang et al., "Detection of SARS-CoV-2 in Different Types of Clinical Specimens," JAMA, vol. 323, no. 18, pp. 1843–1844, 12 2020, doi: 10.1001/jama.2020.3786.

[2] A. Tahamtan and A. Ardebili, "Real-time RT-PCR in COVID-19 detection: issues affecting the results," Expert Rev. Mol. Diagn., pp. 1–2, Apr. 2020, doi: 10.1080/14737159.2020.1757437.

[3] T. Ai et al., "Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases," Radiology, Feb. 2020, doi: 10.1148/radiol.2020200642.

[4] R. M. Pereira, D. Bertolini, L. O. Teixeira, C. N. Silla, and Y. M. G. Costa, "COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios," Comput. Methods Programs Biomed., vol. 194, p. 105532, Oct. 2020, doi: 10.1016/j.cmpb.2020.105532.

[5] W. H. Self, D. M. Courtney, C. D. McNaughton, R. G. Wunderink, and J. A. Kline, "High Discordance of Chest X-ray and CT for Detection of Pulmonary Opacities in ED Patients: Implications for Diagnosing Pneumonia," Am. J. Emerg. Med., vol. 31, no. 2, pp. 401–405, Feb. 2013, doi: 10.1016/j.ajem.2012.08.041.

[6] G. D. Rubin et al., "The Role of Chest Imaging in Patient Management During the COVID-19 Pandemic: A Multinational Consensus Statement From the Fleischner Society," Chest, vol. 158, no. 1, pp. 106–116, 2020, doi: 10.1016/j.chest.2020.04.003.

[7] J.-G. Lee et al., "Deep Learning in Medical Imaging: General Overview," Korean J. Radiol., vol. 18, no. 4, pp. 570–584, Aug. 2017, doi: 10.3348/kjr.2017.18.4.570.

[8] Y. Fang et al., "Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR," Radiology, vol. 296, no. 2, pp. E115–E117, Feb. 2020, doi: 10.1148/radiol.2020200432.

[9] M.-Y. Ng et al., "Imaging Profile of the COVID-19 Infection: Radiologic Findings and Literature Review," Radiol. Cardiothorac. Imaging, vol. 2, no. 1, p. e200034, Feb. 2020, doi: 10.1148/ryct.2020200034.

[10] P. Rajpurkar et al., "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning," ArXiv171105225 Cs Stat, Dec. 2017, Accessed: Sep. 29, 2020. [Online]. Available: http://arxiv.org/abs/1711.05225.

[11] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images," ArXiv200311055 Cs Eess, Mar. 2020, Accessed: Sep. 29, 2020. [Online]. Available: http://arxiv.org/abs/2003.11055.

[12] L. Wang and A. Wong, "COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images," ArXiv200309871 Cs Eess, May 2020, Accessed: Sep. 29, 2020. [Online]. Available: http://arxiv.org/abs/2003.09871.

[13] M. Farooq and A. Hafeez, "COVID-ResNet: A Deep Learning Framework for Screening of COVID19 from Radiographs," ArXiv200314395 Cs Eess, Mar. 2020, Accessed: Sep. 29, 2020. [Online]. Available: http://arxiv.org/abs/2003.14395.

[14] P. K. Sethy and S. K. Behera, "Detection of Coronavirus Disease (COVID-19) Based on Deep Features," Mar. 2020, doi: 10.20944/preprints202003.0300.v1.

[15] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, "Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network," Appl. Intell., vol. 51, no. 2, pp. 854–864, Feb. 2021, doi: 10.1007/s10489-020-01829-7.

[16] C. Zheng et al., "Deep Learning-based Detection for COVID-19 from Chest CT using Weak Label," medRxiv, p. 2020.03.12.20027185, Mar. 2020, doi: 10.1101/2020.03.12.20027185.

[17] Y. Song et al., "Deep learning Enables Accurate Diagnosis of Novel Coronavirus (COVID-19) with CT images," medRxiv, p. 2020.02.23.20026930, Feb. 2020, doi: 10.1101/2020.02.23.20026930.

[18] L. J. M. Kroft, L. van der Velden, I. H. Girón, J. J. H. Roelofs, A. de Roos, and J. Geleijns, "Added Value of Ultra-low-dose Computed Tomography, Dose Equivalent to Chest X-Ray Radiography, for Diagnosing Chest Pathology," J. Thorac. Imaging, vol. 34, no. 3, pp. 179–186, May 2019, doi: 10.1097/RTI.0000000000000404.

[19] N. Chen et al., "Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study," Lancet Lond. Engl., vol. 395, no. 10223, pp. 507–513, Feb. 2020, doi: 10.1016/S0140-6736(20)30211-7.

[20] G. Jain, D. Mittal, D. Thakur, and M. K. Mittal, "A deep learning approach to detect Covid-19 coronavirus with X-Ray images," Biocybern. Biomed. Eng., vol. 40, no. 4, pp. 1391–1405, Oct. 2020, doi: 10.1016/j.bbe.2020.08.008.

[21] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. Jamalipour Soufi, "Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning," Med. Image Anal., vol. 65, p. 101794, Oct. 2020, doi: 10.1016/j.media.2020.101794.

[22] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," Phys. Eng. Sci. Med., vol. 43, no. 2, pp. 635–640, Jun. 2020, doi: 10.1007/s13246-020-00865-4.

[23] H. Panwar, P. K. Gupta, M. K. Siddiqui, R. Morales-Menendez, and V. Singh, "Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet," Chaos Solitons Fractals, vol. 138, p. 109944, Sep. 2020, doi: 10.1016/j.chaos.2020.109944.

[24] J. P. Cohen, ieee8023/covid-chestxray-dataset. 2021.

[25] "Chest X-Ray Images (Pneumonia)." https://kaggle.com/paultimothymooney/chest-xray-pneumonia (accessed Mar. 05, 2021)

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:11:41 UTC from IEEE Xplore. Restrictions apply.

2021 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT) | 978-1-6654-2306-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/CSNT51715.2021.9509616

COVID-19 Detection from Chest X-Rays and CT Scans using Dilated Convolutional Neural Networks

Shailender Kumar, Kumar Abhishek, Kritika, Kashish Singh
Department of Computer Science & Engineering
Delhi Technological University, New Delhi, India
shailenderkumar@dce.ac.in, abhishek750980@gmail.com, kritikakumar26@gmail.com, kashishsingh8700@gmail.com

***Abstract*— WHO has declared "Coronavirus disease 2019" (COVID-19), which is caused by "Severe Acute Respiratory Syndrome Coronavirus 2" (SARS-CoV-2), a worldwide pandemic in March 2020. With its advent in December 2019, it has affected over 86,095,614 people worldwide as of January 4, 2021. Medical workers and researchers are working towards developing a vaccine and improving diagnostic methods for early detection and disease progression monitoring methods. The objective of this study is to provide a robust "Convolutional Neural Network" (CNN) architecture for COVID-19 detection using "Chest X-Rays" (CXR) and Chest CT Scans in order to reduce the response time to diagnose infected patients. We developed deep learning image classification models using Dilated Convolutional Neural Networks as the backend for our model and utilized various fine-tuned pre-trained CNN models as the feature extractor for our model. For both Chest X-Ray and Chest CT, we created datasets by combining various publicly available databases. The Chest X-Ray dataset contains 196 COVID positive frontal CXR images and 196 normal images, and out of two Chest CT datasets, one contains 349 covid and 349 non-covid images and the other contains 1252 covid and 1230 non-covid images. We also utilized transfer learning because of less publicly available data. Image Enhancement Techniques were also used to improve image contrast. The best classification accuracy achieved on Chest X-Ray dataset is 100% and accuracies achieved on the two Chest CT datasets are 91.6% and 98% respectively.**

***Index Terms*— Dilated Convolutional Neural Networks, COVID-19, Transfer Learning, Chest Radiography, Chest Computed Tomography, Deep Learning**

I. INTRODUCTION

COVID-19 is a contagious disease that started to proliferate from Wuhan, China, in December 2019. "It spread across the world and was declared a global pandemic by the World Health Organisation on March, 2020" [21]. As of January 04, 2021, according to Worldometers' website, the number of total COVID-19 global cases has touched 86,095,614 with 23,177,084 active cases [1]. The virus carries symptoms like fever, cough, shortness of breath, and in many cases, it is asymptomatic and shows no sign of illness. Doctors suggest that early diagnosis helps isolate the patient and reduces the chain of spread by a huge factor. COVID-19 has several diagnostic methods like RT-PCR, CXR and chest computed tomography (CCT).

The go-to method preferred for COVID-19 diagnosis is "reverse transcriptase-polymerase chain reaction" (RT-PCR) testing. RT-PCR screening is highly specific to COVID-19, but it also has some limitations like it gives false-negative results [2] and the process is time-consuming and needs specialized test equipment and machines. COVID-19 affects the respiratory system along with visible abnormalities in the lungs. Imaging modes can be utilised at early stages to detect virus and condition of affected organs, but the final report requires accurate manual detection and time. There are chances of human errors, and smaller anomalies might go unnoticed. Therefore, to assist doctors with a second reference for detecting COVID-19 image findings in various imaging modes, we used deep learning methods for faster results and detailed visualisation of medical images. We used Dilated Convolutional Neural Networks for detection of COVID-19 in CT scan and CXR images. As limited data is available in the public domain, we used fine-tuned pre-trained models and transfer learning concepts for feature extraction.

In section II, we briefly explain the advantages of using CNN for image classification and mention critical image findings for CXR and CT scans. We review some former work done on COVID-19 diagnosis which utilised CNN and deep learning. In section III, we propose our approach for COVID-19 detection in both image modalities, which consists of three stages - data collection, data pre-processing and model development. The results obtained in different experiments are mentioned in section IV. In section V, we use Grad-CAM visualisation for explainability of our CNN models. Finally, we conclude our study with a brief summary in section VI.

II. RELATED WORK

Artificial Neural Networks like CNN have proven to be quite beneficial in detecting even the fine details in medical images. Artificial Intelligence and deep learning backed up with transfer learning have proved to be extremely efficient in COVID-19 classification [3]. Dilated Convolutional Neural Networks, in particular, are useful for expanding the

978-1-6654-2306-9/21/$31.00 ©2021 IEEE
DOI: 10.1109/CSNT.2021.65

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:32:19 UTC from IEEE Xplore. Restrictions apply.

receptive field without losing resolution and for faster computation [4]. "In abnormal chest X-ray, consolidation opacities were the most common finding seen in 218 patients (81.3%), followed by reticular interstitial thickening seen in 107 patients (39.9%) and GGO seen in 87 patients (32.5%)" [5]. Xu et al. [6] performed CT scan analysis on 90 patients, and detected minor lung lesions at early stages and other common features of disease progression were adjacent pleural thickening, consolidation, crazy paving pattern, ground-glass opacities, and linear opacities.

Michael et al. [7] used transfer learning and public datasets and stated in their paper that "the selected VGG19 model, which is then extensively tuned with appropriate parameters, performs in considerable levels of COVID-19 detection against pneumonia or normal for all three lung image modes with the precision of up to 86% for X-Ray, 100% for Ultrasound and 84% for CT scans" [7]. Zhang et al. [8] built a customized ResNet network for CXRs COVID-19 images which achieved 96% sensitivity and 70.65% specificity on a dataset of 100 COVID and 1431 non-COVID images. A study [10] performed on ten pre-trained models for classification on private dataset of COVID and Non-COVID CT images along with a comparative analysis claimed that ResNet-101 with an accuracy of 99.51% performed best among others. Other similar studies also provide good evidence for claiming that deep learning and transfer learning can be used to build networks with better accuracy to detect and diagnose COVID-19 features in medical images.

III. METHODOLOGY

A. Data Collection

We collected CXR and CT images of normal, Non-COVID and COVID infected cases from various published datasets available in the public domain. We used Cohen et al. [11] dataset for collecting 196 posteroanterior COVID-19 positive CXR images which are extracted from various websites, online publications and directly from PDF. 196 Normal CXR images are sourced from online available CXR images collection [12] which contains normal and pneumonia infected images. CT scan images are categorized into two classes: COVID and Non-COVID images. Dataset [13] contains a total of 2482 CT scan images classified into two categories: 1252 positive COVID images and 1230 negative COVID images. 349 COVID positive CT scan images and 397 negative images are further extracted from the publicly available dataset [14].

B. Data Pre-processing

Due to the small size of the sampled corpus for both imaging modes, it was a huge challenge to train deep learning models with less training data. To enhance the volume of training-data, we implemented data augmentation technique with various transformations, including vertical flip, horizontal flip, zoom, shear, and rotation. Another major problem with the sampled dataset was that images from different datasets and different imaging modes had intensity and brightness differences, and "medical researchers have indicated that medical image analysis methods are susceptible to variations of image intensities" [15].

TABLE I: Dataset Sources

Image Modality	**Class**	**Datasets**	**Number of Images**
CXR	Covid	Cohen et. al. [11]	196
	Normal	Kermany et. al. [12]	196
CT Scan - I	Covid	Soares et al. [13]	349
	Non-Covid	Soares et al. [13]	349
CT Scan - II	Covid	Yang et al. [14]	1252
	Non-Covid	Yang et al. [14]	1230

To overcome sample bias, we applied histogram equalization to images using the algorithm N-CLAHE [16] to improve image contrast and brightness. "The method both normalizes images and enhances small details, textures and local contrast by first globally normalizing the image histogram followed by application of Contrast Limited Adaptive Histogram Equalization (CLAHE)" [17]. It is evident from Table II that N-CLAHE pre-processing step considerably improved the image brightness, contrast and also enhanced the fine details of the images across all datasets.

C. Model Development

For both imaging modes (X-Ray & CT scan) our proposed architectures are single-column Convolutional Neural Networks consisting of two major components:

1) Front-End of the models built using fine-tuned pre-trained Convolutional Neural Networks like (VGG, ResNet, DenseNet etc).
2) Back-End of the models built using Dilated Convolution Layers along with "Fully-Connected layers" (FC) with a sigmoid activation function at the end of the model to classify input images into different classes.

- **Front-End**: Due to Convolutional neural networks' exceptional performance in image classification, we decided to utilize CNNs as feature extractors for our deep neural network model. Also, data related to COVID patients available publicly is not in abundance which makes the task of COVID-19 detection using CNNs more difficult. So, due to the above-mentioned reasons, we decided to use transfer learning with data augmentation to compensate for less amount of data available. Pre-trained models like VGG [18], ResNet [19], Xception [20] etc. are Convolutional Neural

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:32:19 UTC from IEEE Xplore. Restrictions apply.

Networks trained on the ImageNet database, which comprises of 14 million images belonging to 1000 different classes and their pre-trained weights can be utilized to make custom single-column Convolutional Neural networks which perform really well in cases of fewer data.

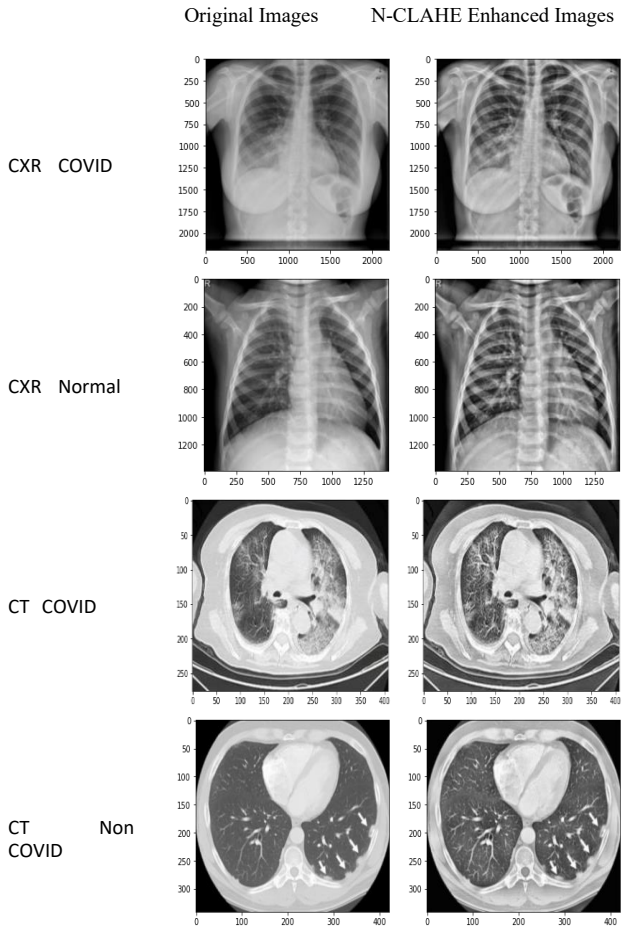


TABLE II: N-CLAHE Enhancement on CXR and CT images

- **Back-End**: After passing through the Front-End of the model, the input image's size has been reduced due to the presence of various pooling layers. Reducing the input image size too much can cause a loss of important information, which will greatly reduce our models' performance. So, we do not want our back-end to decrease the feature maps' size. However, the reduction in the input image size because of pooling layers can be solved by adding Transpose Convolutional layers or Deconvolutional layers [9], but adding these layers also increases the number of model parameters, increasing the model training time. So, instead of pooling and deconvolutional layers, we utilized dilated convolutional layers because these layers drastically reduce the training time as the number of model parameters decreases, and we can also cover a larger receptive field on the input image even with a small kernel size like (3 x 3). Dilated Convolutional Layers also perform exceptionally well in image segmentation tasks.
- **Dilated Convolution**: The most important part of our CNN architectures is the Dilated Convolution layers. A dilated convolution layer is a convolution layer which is applied to the input tensor with defined gaps. A 2D Dilated Convolution operation is described as follows:

$$O(m,n) = \sum_{k=1}^{m} \sum_{l=1}^{n} I(m + r \times k, n + r \times l) f(k,l) \qquad (1)$$

O (m, n) is the output feature map after passing through the dilated convolution layer with a 2D input image as I(m, n) and a 2D kernel f(k, l) where r is the dilation rate. Deconvolutional layers [9] can compensate for the loss of information, but the added complexity and increased training latency may not be suitable for all cases. "This technique enlarges the receptive field without increasing the number of model parameters. In dilated convolution, a small-size kernel with (k x k) filter is enlarged to k + (k - 1) (r - 1) with dilated stride r. Thus, it allows flexible aggregation of the multi-scale contextual information while keeping the same resolution. Examples can be found in Fig. 1 where normal convolution gets (3 x 3) receptive field and two dilated convolutions deliver (5 x 5) and (7 x 7) receptive fields respectively" [4]. In image segmentation tasks, dilated convolution layers have shown excellent results [[22], [23]], and are an overall better alternative to pooling layers. Pooling layers are used in deep learning models to preserve invariance and control overfitting, and also significantly decrease the spatial resolution of the feature maps.

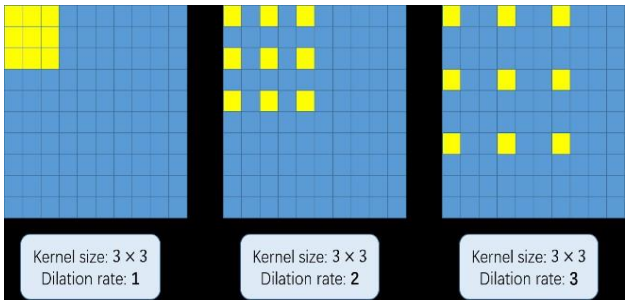


Fig. 1: (3 x 3) filters with various dilation rates, Source: Adapted from [4]

D. Network Configuration

Network Configurations of our CNN architectures for both CXR and CT scan datasets are shown in Table III.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:32:19 UTC from IEEE Xplore. Restrictions apply.

TABLE III: Network Configurations for both CXR and CT scan. The convolutional layers' parameters are denoted as "conv-(number of filters)-(kernel size)-(dilation rate)".

Network Configuration			
A	B	C	D
Input (224 x 224 or 299 x 299 colour image)			
Front-end (fine-tuned CNN model)			
Back-end (four different configuration)			
conv-512-3-1 conv-512-3-1 conv-512-3-1 conv-256-3-1 conv-128-3-1 conv-64-3-1	conv-512-3-2 conv-512-3-2 conv-512-3-2 conv-256-3-1 conv-128-3-1 conv-64-3-1	conv-512-3-3 conv-512-3-3 conv-512-3-3 conv-256-3-3 conv-128-3-3 conv-64-3-3	conv-512-3-4 conv-512-3-4 conv-512-3-4 conv-256-3-4 conv-128-3-4 conv-64-3-4
Global Average Pooling			
FC-256			
FC-256			
Sigmoid			

E. Training Details

We used a straightforward method for training our Convolutional neural networks as end-to-end structures. Firstly, we take the Convolution layers of the pre-trained CNN models without the classifier layers and then fine-tune them according to our dataset. As for the backend, we utilize dilated convolution layers along with FC layers to build the classifier part of our model. During model training, Adam optimizer is used with learning rate (lr = 1e-6) and decay rate = 5e-4. Also, we chose the Binary Cross Entropy as our loss function for the models. The loss function is given as follows:

$$Loss = -\frac{1}{n}\sum_{\theta=1}^{n}(h_\theta.\ln(\widehat{h_\theta}) + (1-h_\theta).\ln(1-\widehat{h_\theta})) \quad (2)$$

where, n = number of values in the model output, $\hat{h}_\theta$ = θ-th value in the model output, and h_θ is the corresponding target value.

IV. Experiments and Results

We evaluated the performance of our models on the i5 7300HQ CPU @ 2.50 GHz, 8GB RAM, GTX 1650 4GB graphics, Windows 10 64-bit PC. Implementation of our CNN models is performed using Python using the Keras library with TensorFlow as backend in the Anaconda environment.

We created & evaluated different single column Convolutional Neural Network models using various pre-trained CNN models at the Front-End with different dilation rates at the Back-End.

A. Chest X-Ray Dataset:

Dataset [11,12] sampled for CXR images contains 196 positive frontal CXR images and 196 normal CXR images. In this experiment, we used various pre-trained CNN models as our front-ends, each with different backends (Table III) using Dilated CNNs. Accuracy achieved by different models are shown in Table IV. It is evident from the results that pre-trained CNNs are performing better with backend B.

TABLE IV: Accuracy achieved on CXR dataset
(A= Accuracy, R= Recall, F1= F1-Score)

Model Front-end		**Model Back-end**			
		A	***B***	***C***	***D***
VGG-16	A	0.99	1.00	0.97	1.00
	R	1.00	1.00	0.98	1.00
	F1	0.99	1.00	0.98	1.00
VGG-19	A	0.98	1.00	1.00	0.96
	R	0.98	1.00	1.00	0.95
	F1	0.98	1.00	1.00	0.96
DenseNet-121	A	0.94	0.99	0.96	0.92
	R	0.93	1.00	0.95	0.93
	F1	0.94	0.99	0.96	0.93
DenseNet-169	A	0.99	0.96	1.00	0.97
	R	1.00	0.95	1.00	0.98
	F1	0.99	0.96	1.00	0.98
Xception	A	0.99	1.00	0.97	0.98
	R	1.00	1.00	0.98	0.98
	F1	0.99	1.00	0.98	0.98
MobileNet	A	0.93	0.95	0.93	0.81
	R	0.95	0.93	0.95	0.91
	F1	0.94	0.95	0.94	0.80

B. Chest CT Datasets:

We performed experiments on two different CT scan datasets [13,14] with the same network configurations as depicted in Table III. The accuracy of the models is shown in Table V and VI.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:32:19 UTC from IEEE Xplore. Restrictions apply.

TABLE V: Accuracy achieved on CT scan dataset [13]
(A= Accuracy, R= Recall, F1 = F1-Score)

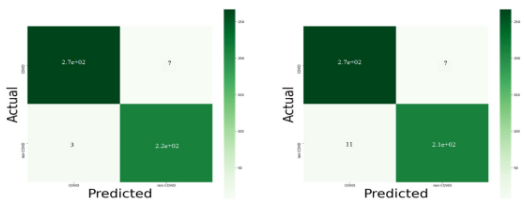
Model Front-end		Model Back-end			
		A	***B***	***C***	***D***
VGG-16	A	0.833	0.906	0.854	0.885
	R	0.915	0.986	0.930	0.958
	F1	0.838	0.910	0.858	0.884
VGG-19	A	0.833	0.854	0.843	0.864
	R	0.915	0.930	0.917	0.920
	F1	0.838	0.858	0.853	0.867
DenseNet-121	A	0.833	0.895	0.864	0.854
	R	0.915	0.985	0.920	0.930
	F1	0.838	0.896	0.867	0.858
MobileNet	A	0.855	0.885	0.864	0.916
	R	0.930	0.958	0.920	0.986
	F1	0.858	0.884	0.867	0.917

Dataset [13] has very limited COVID positive images which restricted the performance of our CNN models. Our models (MobileNet + backend-D (accuracy = 0.916) and VGG-16 + backend-B (accuracy = 0.906)) out-performed the best model, mentioned in dataset study [13], which utilised multi-task learning and contrastive self-supervised learning to achieve the best accuracy of 0.89.

TABLE VI: Accuracy achieved on CT scan dataset [14]
(A= Accuracy, R= Recall, F1 = F1-Score)

Model Front-end		Model Back-end			
		A	***B***	***C***	***D***
VGG-16	A	0.969	0.978	0.974	0.980
	R	0.970	0.971	0.971	0.975
	F1	0.970	0.980	0.976	0.981
VGG-19	A	0.965	0.969	0.978	0.974
	R	0.974	0.970	0.971	0.971
	F1	0.968	0.970	0.980	0.976
DenseNet-121	A	0.951	0.953	0.966	0.956
	R	0.963	0.970	0.974	0.970
	F1	0.956	0.957	0.968	0.959
MobileNet	A	0.943	0.980	0.964	0.964
	R	0.952	0.975	0.970	0.970
	F1	0.951	0.981	0.967	0.967

Out of various pre-trained CNN models, DenseNet-121 [24], VGG-19 [18], VGG-16 [18], and MobileNet [25] performed considerably better on dataset [14] and out of these four models VGG-16 and MobileNet outperformed with the best accuracy = 0.980 and F1-Score = 0.983.

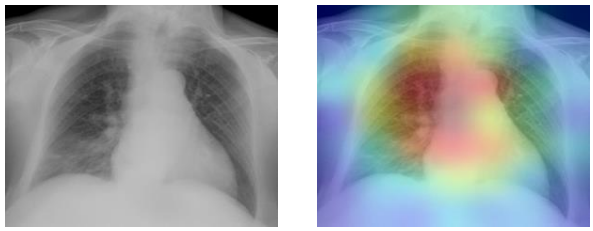


(a) VGG-16 + Backend-D (b) MobileNet + Backend-B

Fig. 2: Confusion Matrix of Models on CT Scan Dataset [14]

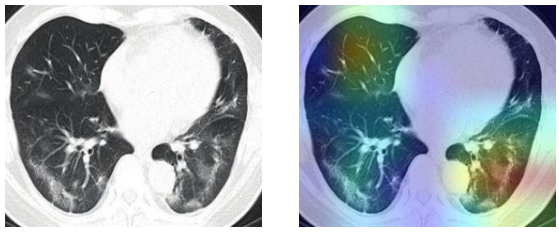
V. VISUALISING RESULTS USING GRAD- CAM

To add more interpretability and transparency to the working of our models, we used Grad-CAM visualisation. It highlighted the parts of CXR and CT scan images that are captured by our models during training. As we can see in fig.3-4, our CNN models localized parts of lungs that showed abnormality, thus visually explaining model's judgement.



(a) Original Image (b) Grad-CAM visualisation

Fig. 3: Grad-CAM visualisation on CXR images



(a) Original Image (b) Grad-CAM visualisation

Fig. 4: Grad-CAM visualisation on CT scan images

VI. CONCLUSIONS

In this study, we experimented on two types of image modalities: CXR and chest CT scans and did binary classification to detect COVID-19. We utilised N-CLAHE enhancement algorithm to improve brightness and contrast of images in both our datasets. Due to less publicly available data of COVID patients, we utilised transfer learning with data augmentation to improve the volume of our training set. The major component of our network architectures, dilated convolution layers, was an over-all better alternative than pooling layers as it reduced training latency and also preserved spatial information in images to enhance the model's performance. For better explainability of our deep CNN models, we used Grad-CAM visualisation on images, and the results showed that our models were able to capture COVID image findings in both CXR and CT images. We are not claiming this model as a primary method for COVID-19 diagnosis, but our models will enable the radiologist to easily spot COVID image findings and reduce response time. As a future direction, we can also use dilated CNNs in the pre-processing stage for lung segmentation which would further enhance the model's performance as the model will focus on critical image features with reduced noise.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:32:19 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] Worldometers Website. [Online]. Available: https://www.worldometers. info/coronavirus/ 2021.

[2] West, C. P., Montori, V. M., & Sampath Kumar, P. (2020). "COVID-19 Testing: The Threat of False-Negative Results," Mayo Clinic proceedings, 95(6), 1127-1129. https://doi.org/10.1016/j.mayocp.2020.04.004

[3] M. E. H. Chowdhury et al., "Can AI Help in Screening Viral and COVID-19 Pneumonia?," IEEE Access, Vol. 8, pp. 132665-132676, 2020, doi: 10.1109/ACCESS.2020.3010287.

[4] Y. Li, X. Zhang and D. Chen, "CSRNet: Dilated Convolutional Neural Networks for Understanding the Highly Congested Scenes," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, 2018, pp. 1091-1100, doi: 10.1109/CVPR.2018.00120.

[5] Yasin, R., Gouda, W.," Chest X-ray findings monitoring COVID-19 disease course and severity," Egypt J Radiol Nucl Med 51, 193 (2020). https://doi.org/10.1186/s43055-020-00296-x

[6] Xu X, Yu C, Qu J, Zhang L, Jiang S, Huang D, Chen B, Zhang Z, Guan W, Ling Z, Jiang R, Hu T, Ding Y, Lin L, Gan Q, Luo L, Tang X, Liu J., "Imaging and clinical features of patients with 2019 novel coronavirus SARS-CoV-2," Eur J Nucl Med Mol Imaging, 47(5):1275-1280, 2020.

[7] M. J. Horry et al., "COVID-19 Detection Through Transfer Learning Using Multimodal Imaging Data," in IEEE Access, vol. 8, pp. 149808-149824, 2020, doi: 10.1109/ACCESS.2020.3016780.

[8] J. Zhang, Y. Xie, Y. Li, C. Shen, and Y. Xia, "Covid-19 screening on chest x-ray images using deep learning based anomaly detection," arXiv preprint arXiv:2003.12338, 2020.

[9] M. D. Zeiler, D. Krishnan, G. W. Taylor and R. Fergus, "Deconvolutional networks," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, 2010, pp. 2528-2535, doi: 10.1109/CVPR.2010.5539957.

[10] Ardakani AA, Kanafi AR, Acharya UR, Khadem N, Mohammadi A.,"Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks, " Comput Biol Med. 2020;121:103795. doi:10.1016/j.compbiomed.2020.103795.

[11] J. P. Cohen, P. Morrison, L. Dao, K. Roth, T. Q. Duong, and M. Ghassemi, "COVID-19 image data collection: Prospective predictions are the future," 2020, arXiv:2006.11988. [Online]. Available: http://arxiv.org/abs/2006.11988

[12] Kermany, Daniel; Zhang, Kang; Goldbaum, Michael (2018), "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification," Mendeley Data, V2, doi: 10.17632/rscbjbr9sj.2.

[13] Soares, Eduardo, Angelov, Plamen, Biaso, Sarah, Higa Froes, Michele, and Kanda Abe, Daniel. "SARS-CoV-2 CT-scan dataset: A large dataset of real patients CT scans for SARS-CoV-2 identification." medRxiv (2020). doi: https://doi.org/10.1101/2020.04.24.20078584.

[14] Yang, X., Zhao, J., Zhang, Y., He, X., & Xie, P. (2020),"COVID-CT-Dataset: A CT Scan Dataset about COVID-19," ArXiv, abs/2003.13865.

[15] S. Song, Y. Zheng, and Y. He, "A review of methods for bias correction in medical images," Biomed. Eng. Rev., vol. 1, no. 1, 2017, doi: 10.18103/bme.v3i1.1550.

[16] K. Koonsanit, S. Thongvigitmanee, N. Pongnapang, and P. Thajchayapong, "Image enhancement on digital X-ray images using NCLAHE," in Proc. 10th Biomed. Eng. Int. Conf. (BMEiCON), Hokkaido, Japan, Aug. 2017, pp. 14, doi: 10.1109/BMEiCON.2017.8229130.

[17] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in Graphics Gems IV. New York, NY, USA: Academic, 1994.

[18] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, arXiv:1409.1556. [Online]. Available: http://arxiv.org/abs/1409.1556.

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, Jun. 2016, pp. 770–778, doi:10.1109/CVPR.2016.90.

[20] F. Chollet, "Xception: Deep learning with depth wise separable convolutions," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Honolulu, HI, USA, Jul. 2017, pp. 1800–1807, doi:10.1109/CVPR.2017.195.

[21] Who COVID-19 Situation Reports. [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports, 2020.

[22] Fisher Yu and Vladlen Koltun.,"Multi-scale context aggregation by dilated convolutions," In ICLR, 2016.

[23] L. Chen, G. Papandreou, I. Kokkinos, K. Murphy and A. L. Yuille, "Deep Lab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 4, pp. 834-848, 1 April 2018, doi: 10.1109/TPAMI.2017.2699184.

[24] G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 2261-2269, doi: 10.1109/CVPR.2017.243.

[25] Howard, Andrew G., et al.,"Mobilenets: Efficient convolutional neural networks for mobile vision applications," arXiv preprint arXiv:1704.04861 (2017).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:32:19 UTC from IEEE Xplore. Restrictions apply.

10th IEEE International Conference on Communication Systems and Network Technologies

2021 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT) | 978-1-6654-2306-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/CSNT51715.2021.9509636

COVID-19 Tweet Analysis using Hybrid Keyword Extraction Approach

Sakshi Vatsa, Surbhi Mathur, Mansi Garg, Rajni Jindal
Dept. of computer engineering,
Delhi Technological University, Delhi, India
Email: vatsasakshi@gmail.com, surbhimathur1210@gmail.com, gargmansi105@gmail.com, rajnijindal@dce.ac.in

Abstract:- Aim of this paper is to identify the key issues discussed among the people on Twitter using machine learning and NLP techniques regarding COVID-19. One of the most important way to produce these insights is Automatic Keyword Extraction. It is a method to obtain the most important words from the text by summarizing thus providing an insight into the whole context. Text Summarization is the process to condense the content without the loss of significant data. This paper applies a hybrid model of graph-based and topic modeling approaches to extract keywords from a large dataset of approximately 1 million tweets.

Keywords: covid-19, Automatic Keyword Extraction, Graph-based, Clustering, Community, Topic modeling, tweet analysis

I. INTRODUCTION

Since the start of 2020, millions of people worldwide have been found positive with the CoronaVirus. We need to understand the significant problems faced by the people due to this fast transmission of COVID-19. In this paper, we use covid-19 related tweets to perform this analysis. Keywords highlight the immediate context in any documentation. The various applications are text summarization, information retrieval, word cloud generation, and many more. The current approaches for keyword extraction can be divided into two main categories, Supervised and Unsupervised learning. The Supervised learning approach is a binary classification problem that classifies each word as a keyword or not. The supervised learning approaches use classifiers like SVM, Neural Networks, Decision trees, or logistic regression for classification. However, they require labeled input data for training which is highly infeasible for a large data corpus. The unsupervised learning-based Keyword Extraction approaches can be broadly divided into three main categories: 1.) Statistical, 2.) Graph-based and 3.) Topic modeling approaches. Statistical approaches use mathematical factors like word frequency, positional features, POS tags, which help understand the context's mathematical distribution. The whole textual data is viewed as a combination of various topics in the topic modeling approach. This approach works quite well for long documents, but extending to small documents does not provide satisfactory results. The third approach used for Keyword Extraction is a graph-based model representing a weighted graph that also helps incorporate the semantic relationships between words. Though graph-based approaches have proved to work efficiently, they are incapable of covering various topics in the text. To handle this drawback, we proposed a hybrid of topic modeling and a graph-based approach. The main contributions of this paper are:

i) Our approach is an amalgamation of 3 approaches - statistical, graphical, and topic modeling to use the benefits of all the approaches and improve the quality of results.
ii) We have given importance to Hashtags in the tweets, which play an essential role in keyword extraction but have been highly neglected till now.
iii) We have separately treated unigrams and n-grams to distinguish the significance of n-grams based on their frequency comparisons.

II. LITERATURE REVIEW

Xue, Jia, et al. [10] analyze the Twitter data related to Covid 19 between March 7 to April 21, 2020, using LDA, sentiment analysis, and Qualitative Content Analysis to get the tweets' insight. Samuel, Jim, et al. [11] demonstrate insights into the progress of fear-sentiment over time as COVID-19 approached peak levels in the United States, using descriptive textual analytics. [1-3] demonstrates a variety of approaches for keyword extraction methods. [4-5] proposes graph-based keyword extraction methods widely experimenting on different types of centrality measures like degree, closeness, coreness, and neighborhood. Biswas et al. [6] the paper proposed a graph-based approach for keyword extraction on Twitter data. For Candidate Keyword Selection, this paper uses the standard techniques of stopword removal, length filtering, and removal of the unimportant token by using a parameter named Average Occurrence Frequency. We observed that the removal of hashtags gives a drawback as they provide to a great extent a tweet's gist by the user itself and hence handle it in our model. Mao, Xiangke, et al. [7] provides a comprehensive graph-based approach. In the Candidate Keyword Selection, step the paper uses pos-tag filtering by considering only nouns and adjectives. The paper inspired our model to perform pos-tag ranking such that keywords belonging to the uncommon pos tag categories can contribute to the results. For graph construction, the paper

978-1-6654-2306-9/21/$31.00 ©2021 IEEE
DOI: 10.1109/CSNT.2021.25

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:35:30 UTC from IEEE Xplore. Restrictions apply.

focuses on using word co-occurrence and semantic scores. Wang, Hongbin, et al. [8] again present a graphical model that mainly focuses on enhancing the semantic relationship between words by calculating the *word attraction score*. It also factors in positional statistics suggesting a word's position plays a crucial factor in deciding its importance. Yang, Fan et al. [9] also represent a graph-based keyword extraction approach that not only factors in finding important keywords but also tries to find essential relationships between words and sentences and extract keyphrases from the text document. The semantic relationship used in our paper is inspired by [7-9]. We use word co-occurrence for constructing edges. Considering an edge between two words if they occur in the same tweet more than a threshold number of times ensures semanticity. The different variations of distance measure used in [7-9] do not work well with small-sized tweets, and thus our model provides a solution for this. Jindal et al. [12] worked on software bug reports. He proposed a method to summarize the contents using Rapid Automatic Keyword Extraction (RAKE) to score the sentences and the Fuzzy C-Means Clustering to reduce redundancy. And then, both are integrated by predefined rules and followed by Hierarchical Clustering to reduce redundancy. In [15] authors presented improvement in Naïve basian keyword extraction, which could be used for any web state and content mining. Wang, Yuan, et al. [16] proposed a topic model developed to handle tweets using Hashtag Graph-based Topic Model (HGTM). This model utilizes hashtags in the graph by establishing semantic relations based on hashtags.

III. METHODOLOGY

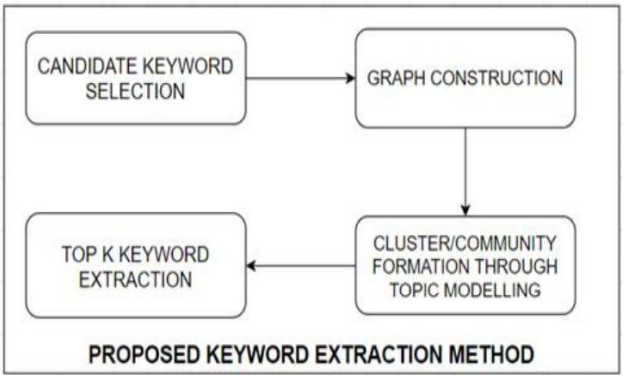


Fig.1 Proposed Framework

A. Data Collection

Twitter data is extracted by the TWEEPY library present in python. We used a predefined set of search words to extract data. The data is collected from the API using the search words coronavirus and covid-19. It comprises around 1 million tweets tweeted in the first week of October containing the expression coronavirus or covid-19.

B. Data Preprocessing

The necessary preprocessing steps needed to extract keywords like - Removing Usernames, URLs, Remove punctuations and numbers, Tokenization, Correct Spellings, Lemmatization/Stemming were followed.

The steps also contain forming a separate list of:

1. *N-grams*: The idea behind this is that their statistical values are low compared to their unigram selves, thus reducing their significance.
2. *Hashtags*: Hashtags provide essential insight into a tweet.

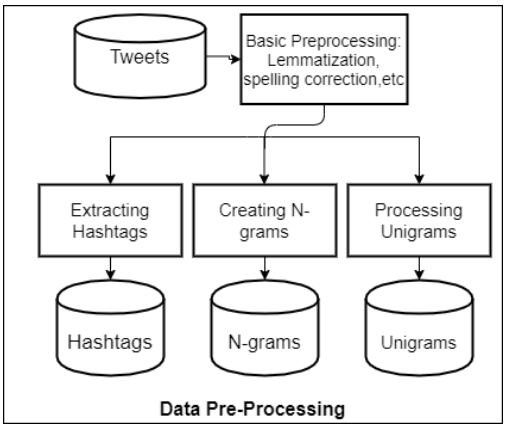


Fig.2 Data Pre Processing Steps

C. Candidate Keyword Selection

Each approach adds a fraction of the final score allotted to each word.

1. POS TAG scoring
2. Modified TF – IDF approach represented by the variable "*tf_idf_score*."
3. *tf_idf_score* = (Modified Term Frequency * Inverse Document Frequency)
4. Modified Term Frequency = Maximum frequency of a word in a tweet.

Word scores are allotted as the weighted sum of the factors mentioned above:

WORD SCORE = (x1) * (pos_tag_value) + (x2) * (tf_idf_score) *where x1, x2 are constants*.

This score is calculated for hashtags, unigrams, N-grams separately. Then the top threshold% of the words are selected to form the candidate keywords.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:35:30 UTC from IEEE Xplore. Restrictions apply.

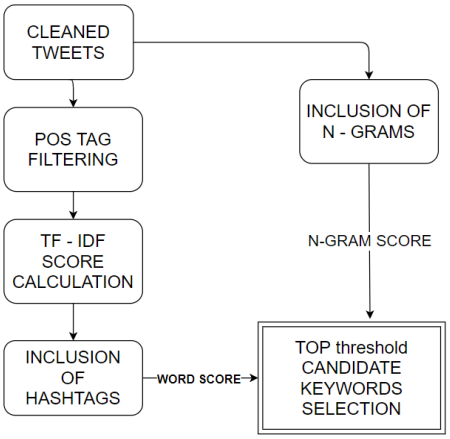


Fig. Candidate Keyword Selection Process

D. Graph Generation

Three main components are :

i. *VERTEX CREATION*: Candidate keywords form the vertices.

ii. *EDGE CREATION:* Edges are formed based on two factors: Firstly, co-occurrence frequency. Edge is created between two vertices if they occur consecutively in the text or if the number of times they appear in the same tweet exceeds a threshold. The latter improves the semantic relationship between the vertices.

iii. *EDGE WEIGHT ALLOCATION:* Currently, the edge weight given to two consecutive vertices is equal to their co-occurrence frequency called *"CO-OCCURRENCE VALUE"* and two non – consecutive vertices is equal to the number of times they occur in the same tweet together, also referred to as "*SEMANTIC SCORE*."

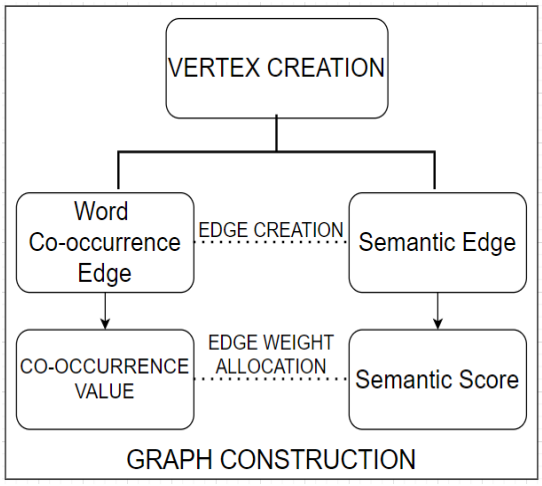


Fig. 4 Construction of Graph

E. Cluster formation

The words are converted to numeric value using the Word2vec library. We are using K-Means Clustering to form the clusters. The algorithm is: 1 Assign each token a cluster. 2, Calculate the Center of the cluster 3, Update all the tokens according to the new center calculated using:

$v_i = (1/c_i) \sum_{j=1}^{c_i} x_{i_j}$, where *'c_i'* represents the number of data points in i^{th}cluster.

F. TOP - K keywords extraction:

Once each vertex is assigned to a community, a vertex score is calculated for each vertex, including the summation of the above-defined "WORD SCORE" and a few other factors. The factors are as follows:

i. *Community_outdegree_value*: This is a scaled value for the number of vertices a node is connected to outside its community.

ii. *Community_intra_degree_value*: This is a scaled value calculated by checking the number of nodes connected inside its community.

iii. *Neighbouring_node_value:* Summation of the score of neighboring nodes divided by the total number of neighboring nodes.

Adding all the above four values gives us the importance of each vertex defined by "*Node_importance."*

*Node_importance=x1*Community_outdegree_value + x2 * Community_intra_degree_value + x3 * Neighbouring_node_value + x4 * word score , where x1,x2,x3,x4 are constants.*

For each community, the number of keywords to be extracted from it is calculated by:

K_i= *Size of the community/ Total size of the graph* such that, $\sum_{i=0}^{n} K_i = K$

The resultant list contains the TOP K selected keywords.

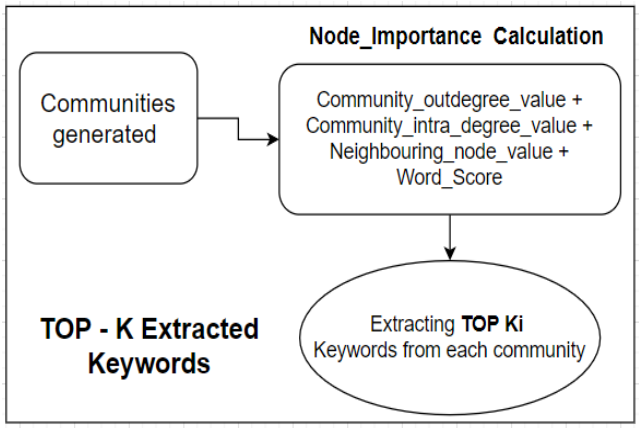


Fig. 5 Top k Resultant Keyword Extraction

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:35:30 UTC from IEEE Xplore. Restrictions apply.

IV. RESULTS

We have represented the keywords in word clouds separately for Hashtags, N-grams, and Unigrams. The following topics are obtained from the above-combined result, as shown in Table 1.

Table 1: Some Community-wise extracted keywords

S. No	KEYWORDS	Topic
1	People, trump, virus, American president, fake_vote, election, control, economy, stay_home, stay_safe, second wave, american dead, kill_american, remove_mask, prevent_spread, total_death	American news
2	School, close, student, restriction, university, teacher, become_infect, mask social distance, super_spreader_event, self_isolation	Education
3	Health, care, help, family, social, distancing, report, service, emergency, wear_mask, public_health	Healthcare
4	Covid, country, world, news, recovery, outbreak, Population, situation,	Worldwide outbreak
5	Government, support, business, challenge, national, policy, fund, economy	Government
6	Wuhan, transparency, real_emerngency_case, circumstances, defence,carrier	China
7	Women, kid, suffer ,disease, condition, parent	Affected people
8	Test ,vaccine, safety, guideline, transmission, prevention, deadly_virus, health_ministry, survival_life, survival_rate, economic_recovery	Vaccination
9	Loneliness, meditation, harmless, inspection, despise, isolate_lonely, risk_factor, mental_health	Mental health
10	Death, report, infection, confirm, hospitalisation, positive, fatality, new_case, active_case, recovery_rate	Statistics
11	White_house , joe_biden, debate, campaign, official, donald_trump, american_die, return_white_house, president_donald	Trump vs Biden

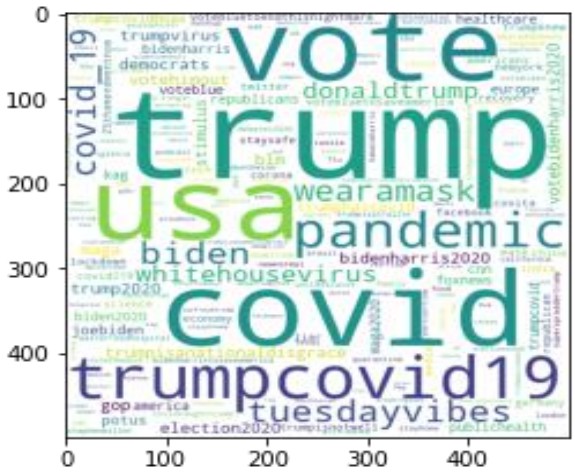


Fig 6: Word cloud of popular hashtags

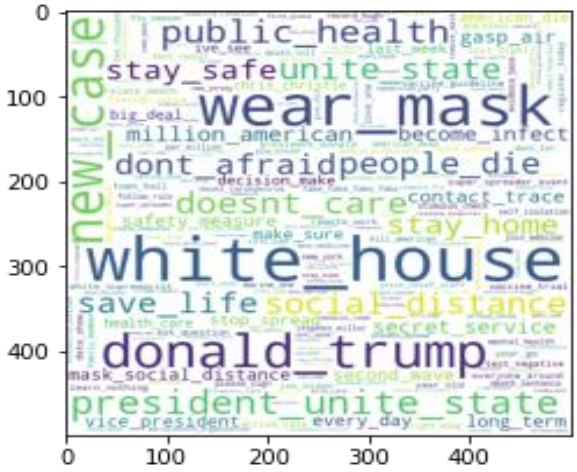


Fig 7: Word cloud of popular n-grams

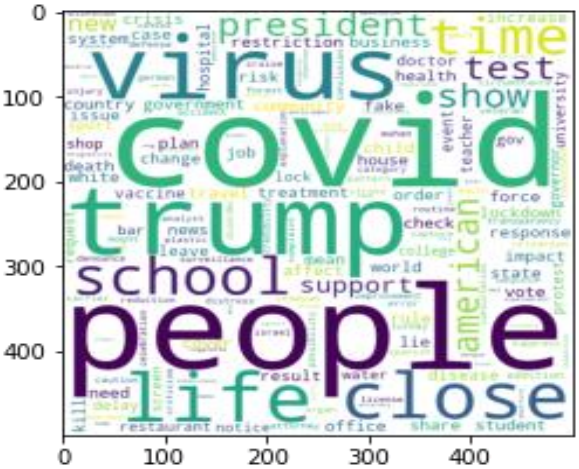


Fig 8: Word cloud of popular unigrams

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:35:30 UTC from IEEE Xplore. Restrictions apply.

V. EXPERIMENTAL ANALYSIS

A pre-labeled dataset from the National Library of Medicine named NLM500 was used to calculate performance metrics. The dataset is described in Table 2.

Table 2: NLM dataset statistics

Number of documents	500
Average document length, in words.	4854
Average keywords assigned per document.	27
Average keyphrases assigned per document.	14
Average percentage of keyphrases present in the text	71.35
Description	Full papers from *PubMed* database

Table 3: Performance metrics of different models on NLM500 dataset.

Model	Precision	Recall	F1 score
RAKE	0.0499	0.0515	0.0506
TextRank	0.1243	0.1142	0.1190
TF-IDF	0.1786	0.1784	0.1739
Proposed Model	**0.1623**	**0.2798**	**0.2051**

As clearly indicated in Table 3, our proposed model outperforms all other models when compared by F1 score and recall. It shows a significant increase of 0.03 in F1 score and an increase of 0.10 in recall.

VI. DISCUSSION

The approach tries to eliminate the individual disadvantages of the methods. Statistical models are known to ignore semantic relationships. The graphical models limit the range of keywords to particular high-frequency topics. Thus the use of statistical models is limited to the Candidate Keyword selection step. The graphical model is used with a clustering approach to cover a wide - variety of topics. N-grams add meaning to otherwise individually irrelevant words. Hashtags help to find user-generated keywords. Thus, processing them separately provides more insight.

VII. CONCLUSION

This paper gives HashTags and N-grams particular importance by using them separately as the dataset, which has shown better results than previous papers. User-generated keywords inspired us to use hashtags separately. Along with it, we have introduced a Novel hybrid model of all the three techniques, i.e., Statistical, Graphical, and Topic modeling, which handles each approach's drawbacks and dramatically improves the current approaches' quality. The major drawback of this approach is that it would leave out many tweets and thus valuable data. During the data retrieval step, the Tweepy library used generated keywords based on specific search words provided. The current model's future scope would include the use of various well-known methods like Text - Rank, Page - Rank to enhance the model's current performance.

REFERENCES

1. S. Duari and V. Bhatnagar, ''SCAKE: Semantic connectivity aware keyword extraction,'' Inf. Sci., vol. 477, pp. 100–117, Mar. 2019.
2. S. Beliga, A. Meštrović, and S. Martinčić-Ipšić, ''An overview of graph based keyword extraction methods and approaches,'' J. Inf. Organizational Sci., vol. 39, no. 1, pp. 1–20, 2015.
3. E. Papagiannopoulou and G. Tsoumakas, A Review of Keyphrase Extraction. Hoboken, NJ, USA: Wiley, 2019
4. Lahiri, S., Choudhury, S. R., & Caragea, C. (2014). Keyword and keyphrase extraction using centrality measures on collocation networks. arXiv:1401.6571 [cs.CL].
5. Abilhoa, W. D., & Castro, L. N. de. (2014). A keyword extraction method from twitter messages represented as graphs. Applied Mathematics and Computation, 240, 308–325.
6. Biswas, Saroj Kr, Monali Bordoloi, and Jacob Shreya. "A graph based keyword extraction model using collective node weight." *Expert Systems with Applications* 97 (2018): 51-59.
7. Mao, Xiangke, et al. "Automatic Keywords Extraction Based on Co-Occurrence and Semantic Relationships Between Words." *IEEE Access* 8 (2020): 117528-117538.
8. Wang, Hongbin, et al. "Unsupervised Keyword Extraction Methods Based on a Word Graph Network." *International Journal of Ambient Computing and Intelligence (IJACI)* 11.2 (2020): 68-79.
9. Yang, Fan, Yue-Sheng Zhu, and Yu-Jia Ma. "WS-rank: Bringing sentences into graph for keyword extraction." *Asia-Pacific Web Conference*. Springer, Cham, 2016.
10. Xue, Jia, et al. "Twitter discussions and emotions about covid-19 pandemic: a machine learning approach." *arXiv preprint arXiv:2005.12830* (2020).
11. Samuel, Jim, et al. "Covid-19 public sentiment insights and machine learning for tweets classification." *Information* 11.6 (2020): 314.
12. Jindal, Shubhra Goyal, and Arvinder Kaur. "Automatic Keyword and Sentence-Based Text Summarization for Software Bug Reports." *IEEE Access* 8 (2020): 65352-65370.
13. Dutta, Soumi, et al. "Ensemble algorithms for microblog summarization." *IEEE Intelligent Systems* 33.3 (2018): 4-14.
14. T.-Y. Liu, "Learning to rank for information retrieval," Foundations and Trends in Information Retrieval, vol. 3, no. 3, 2009, pp. 225–331.
15. G.S. Tomar, S. Verma & Ashish Jha; "Web Page Classification using Modified naïve Baysian Approach", IEEE TENCON-2006, pp 1-4, 14-17 Nov 2006.
16. Wang, Yuan, et al. "Hashtag graph based topic model for tweet mining." *2014 IEEE International Conference on Data Mining*. IEEE, 2014.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:35:30 UTC from IEEE Xplore. Restrictions apply.

Aalto University

OPEN ACCESS

This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Bhardwaj, Shanu; Kesari, Kavindra Kumar; Rachamalla, Mahesh; Mani, Shalini; Ashraf, Ghulam Md; Jha, Saurabh Kumar; Kumar, Pravir; Ambasta, Rashmi K.; Dureja, Harish; Devkota, Hari Prasad; Gupta, Gaurav; Chellappan, Dinesh Kumar; Singh, Sachin Kumar; Dua, Kamal; Ruokolainen, Janne; Kamal, Mohammad Amjad; Ojha, Shreesh; Jha, Niraj Kumar

CRISPR/Cas9 gene editing

DOI:
10.1016/j.jare.2021.07.001

E-pub ahead of print: 01/01/2021

Document Version
Publisher's PDF, also known as Version of record

Published under the following license:
CC BY-NC-ND

Please cite the original version:
Bhardwaj, S., Kesari, K. K., Rachamalla, M., Mani, S., Ashraf, G. M., Jha, S. K., Kumar, P., Ambasta, R. K., Dureja, H., Devkota, H. P., Gupta, G., Chellappan, D. K., Singh, S. K., Dua, K., Ruokolainen, J., Kamal, M. A., Ojha, S., & Jha, N. K. (2021). CRISPR/Cas9 gene editing: New hope for Alzheimer's disease therapeutics. *Journal of Advanced Research*. https://doi.org/10.1016/j.jare.2021.07.001

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

Journal of Advanced Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Advanced Research

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

CRISPR/Cas9 gene editing: New hope for Alzheimer's disease therapeutics

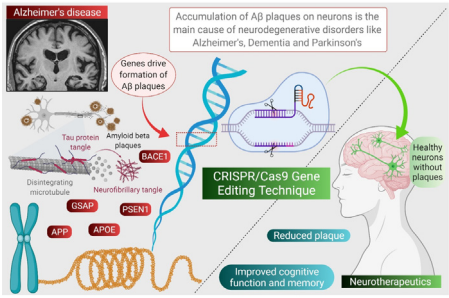
Shanu Bhardwaj [a,1], Kavindra Kumar Kesari [b], Mahesh Rachamalla [c], Shalini Mani [d], Ghulam Md. Ashraf [e,f], Saurabh Kumar Jha [g], Pravir Kumar [h], Rashmi K. Ambasta [h], Harish Dureja [i], Hari Prasad Devkota [j], Gaurav Gupta [k], Dinesh Kumar Chellappan [l], Sachin Kumar Singh [m], Kamal Dua [n,o], Janne Ruokolainen [b], Mohammad Amjad Kamal [p,q,r], Shreesh Ojha [s,*], Niraj Kumar Jha [g,*,1]

[a] Department of Biotechnology, Delhi Technological University (Formerly DCE), Delhi 110042, India
[b] Department of Applied Physics, School of Science, Aalto University, 00076 Espoo, Finland
[c] Department of Biology, University of Saskatchewan, 112 Science Place, Saskatoon, SK S7N 5E2, Canada
[d] Department of Biotechnology, Centre for Emerging Disease, Jaypee Institute of Information Technology, Noida, India
[e] Pre-Clinical Research Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia
[f] Department of Medical Laboratory Technology, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah, Saudi Arabia
[g] Department of Biotechnology, School of Engineering & Technology (SET), Sharda University, Greater Noida 201310, India
[h] Molecular Neuroscience and Functional Genomics Laboratory, Department of Biotechnology, Delhi Technological University (Formerly DCE), Delhi 110042, India
[i] Faculty of Pharmaceutical Sciences, Maharshi Dayanand University, Rohtak, India
[j] Graduate School of Pharmaceutical Sciences, Kumamoto University, 5-1 Oe-honmachi, Kumamoto 862-0973, Japan
[k] School of Pharmacy, Suresh Gyan Vihar University, Mahal road, Jagatpura, Jaipur, India
[l] Department of Life Sciences, School of Pharmacy, International Medical University, Bukit Jalil, Kuala Lumpur 57000, Malaysia
[m] School of Pharmaceutical Sciences, Lovely Professional University, Phagwara-144411, Punjab, India
[n] Discipline of Pharmacy, Graduate School of Health, University of Technology Sydney, Sydney, NSW 2007, Australia
[o] School of Pharmaceutical Sciences, Shoolini University of Biotechnology and Management Sciences, Solan 173229, Himachal Pradesh, India
[p] West China School of Nursing / Institutes for Systems Genetics, Frontiers Science Center for Disease-related Molecular Network, West China Hospital, Sichuan University, Chengdu 610041, Sichuan, China
[q] King Fahd Medical Research Center, King Abdulaziz University, P. O. Box 80216, Jeddah 21589, Saudi Arabia
[r] Enzymoics, NSW 2770; Novel Global Community Educational Foundation, Australia
[s] Department of Pharmacology and Therapeutics, College of Medicine and Health Sciences, United Arab Emirates University, Al Ain, 17666, United Arab Emirates

HIGHLIGHTS

- Mutations in APP, PSEN1 and PSEN2 are known factors for AD pathobiology.
- CRISPR/Cas9 genome editing approach hold promises in AD management.
- CRISPR/Cas9 is utilized to help correct anomalous genetic functions.
- Off-target mutations may impair the functionality of edited cells.
- Non-viral vectors show better efficacy and safety than viral vectors.

GRAPHICAL ABSTRACT



ARTICLE INFO

ABSTRACT

Peer review under responsibility of Cairo University.

* Corresponding authors at: Department of Biotechnology, School of Engineering & Technology (SET) Sharda University, Knowledge Park III, Greater Noida, Uttar Pradesh 201310, India (N.K. Jha). Department of Pharmacology and Therapeutics, College of Medicine and Health Sciences, UAE University, PO Box – 17666, Al Ain, United Arab Emirates (S. Ojha).

E-mail addresses: shreeshojha@uaeu.ac.ae (S. Ojha), nirajkumarjha2011@gmail.com, niraj.jha@sharda.ac.in (N.K. Jha).

[1] These authors contributed equally to this work

https://doi.org/10.1016/j.jare.2021.07.001
2090-1232/© 2021 The Authors. Published by Elsevier B.V. on behalf of Cairo University.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Please cite this article as: S. Bhardwaj, Kavindra Kumar Kesari, M. Rachamalla et al., CRISPR/Cas9 gene editing: New hope for Alzheimer's disease therapeutics, Journal of Advanced Research, https://doi.org/10.1016/j.jare.2021.07.001

Article history:
Received 17 May 2021
Revised 27 June 2021
Accepted 2 July 2021
Available online xxxx

Keywords:
Alzheimer's disease
Neurodegeneration
CRISPR/Cas9
Gene editing
Therapeutics
Presenilin
APP

Background: Alzheimer's disease (AD) is an insidious, irreversible, and progressive neurodegenerative health condition manifesting as cognitive deficits and amyloid beta (Aβ) plaques and neurofibrillary tangles. Approximately 50 million individuals are affected by AD, and the number is rapidly increasing globally. This review explores the role of CRISPR/Cas9 gene editing in the management of AD and its clinical manifestations.

Aim of Review: This review aims to provide a deep insight into the recent progress in CRISPR/Cas9-mediated genome editing and its use against neurodegenerative disorders, specifically AD. However, we have referred to its use against parkinsons's disease (PD), Huntington's disease (HD), and other human diseases, as is one of the most promising and emerging technologies for disease treatment.

Key Scientific Concepts of Review: The pathophysiology of AD is known to be linked with gene mutations, that is, presenilin (PSEN) and amyloid beta precursor protein (APP). However, clinical trials focused at the genetic level could not meet the desired efficiency. The CRISPR/Cas9 genome editing tool is one of the most powerful technologies for correcting inconsistent genetic signatures and now extensively used for AD management. It has significant potential for the correction of undesired gene mutations associated with AD. This technology has allowed the development of empirical AD models, therapeutic lines, and diagnostic approaches for better understanding the nervous system, from *in vitro* to *in vivo* models.
© 2021 The Authors. Published by Elsevier B.V. on behalf of Cairo University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Alzheimer's disease (AD) is a major worldwide health concern with enormous social and economic impact [1]. The slow decline in cognitive function and irreversible neuronal loss are the primary etiological manifestations of the disease. AD is a chronic neurodegenerative condition in which cognition and memory formation gradually deteriorate due to an irreversible neuron loss. It is characterized by the formation and accumulation of amyloid-beta 42 (Aβ42) and phosphorylated Tau, along with excessive glial cell activation. Further, impaired synaptic function and insufficient neurotrophin signaling are important characteristics of AD. Primary symptoms include memory loss, apathy, depression, and irritability [2]. Despite substantial researches, the etiology, pathophysiology, and mechanisms of both cognitive impairment and synaptic dysfunction are not well characterized. In addition, the available therapeutic options are merely symptomatic and supportive, with side effects such as confusion, dizziness, depression, constipation, and diarrhea [3]. Progress towards effective disease-modifying therapies has proven challenging despite the wealth of knowledge on the molecular underpinnings of AD. For instance, several clinical trials have failed to meet efficacy standards against Aβ production, accumulation, and toxicity. This put questions on the amyloid beta hypothesis and advocates for additional treatment strategies. One of the very important and recently invented strategies, clustered regularly interspaced short palindromic repeats/CRISPR-associated proteins 9 system (CRISPR/Cas9) gene editing, attracted attention for possible benefits in the management and treatment of AD. This emerging technology is relatively straightforward, inexpensive, and precise, which has led to an increased interest in this technique for neurodegenerative diseases (NDDs). This can be utilized as a direct treatment approach or may assist in establishing better animal models that faithfully mimic human NDDs. Though, this technique has shown promise in other NDDs such as Huntington's disease (HD) and Parkinson's disease (PD). But, the potential of this technology in the management of AD has not been thoroughly discussed or documented elsewhere. Therefore, the purpose of this review was to examine the potential utility of CRISPR/Cas9 as a treatment option for AD by targeting specific genes, including those that cause early onset AD, as well as those that are significant risk factors for late-onset AD such as the apolipoprotein E4 (APOE4) gene. This review also discusses various delivery systems that help in the proper and targeted delivery of CRISPR/Cas9 cargo in cells.

Alzheimer's disease at a glance

AD is a ubiquitous form of dementia that affects the health of millions of individuals globally. Despite being known for over a century, numerous questions about its pathophysiology are yet to be answered. The standard clinical characteristics of AD are displayed by diminished cognitive abilities, including memory, recognition, judgment, and problem solving [4-6]. Studies on the AD brain revealed neuropathological changes that represent the hallmarks of the condition, such as intracellular neurofibrillary tangles (NFTs) comprising hyperphosphorylated Tau and accumulation of extracellular Aβ plaques [7-9]. In early onset AD, symptoms are exhibited in individuals aged between 30 and 65 years and is primarily genetic, as observed in > 92% cases reported [6], whereas in late onset AD, symptoms start after 65 years of age. As per a 2019 report, >5.8 million individuals in the United States (USA) alone have been diagnosed with AD, out of which 45% of subjects are reported in the 75–84 years group [10]. The constantly increasing numbers of affected patients leads to projections of approximately 14 million in the USA by 2050 [11].

Mostly, the disease is believed to be due to external factors other than genetic predisposition. The amyloid hypothesis explains Aβ formation and aggregation in the brain. This hypothesis states that amyloid beta precursor protein (APP) undergoes proteolysis due to the concerted activities of α-, β-, and γ-secretases. An increase in β-secretase 1 (BACE1) activity is responsible for AD owing to aggregation of Aβ monomers into oligomers, and consequently the generation and deposition of Aβ plaques. Furthermore, BACE1-cleaved APP leads to the formation of a C99 fragment further cleaved by the γ-secretase to produce Aβ monomers Aβ40 and Aβ42. Interestingly, the α-secretase can also cleave APP at various sites, thus curtailing Aβ monomer generation (Fig. 1**A**) [12].

The amyloid hypothesis explains Aβ formation; similarly, NFT formation in AD brains is revealed by the popular Tau hypothesis. Tau, a well-known microtubule associated protein, plays a pivotal role in the formation and stabilization of the microtubule cytoskeleton [13]. It has been reported that out of six Tau isoforms, 3R and 4R are prominent in adult human axons. Multiple phosphates and kinases have Tau as target. In the AD brain, Tau isoforms 3R and 4R can accumulate in hyperphosphorylated form and cause NFT formation in neuronal tissues, if present in neuronal cell axons and bodies, leading to Tau pathology. As recently suggested, Tau oligomers can be neuropathology-mediating microstructures or potential molecular initiators in AD (Fig. 1**B**)

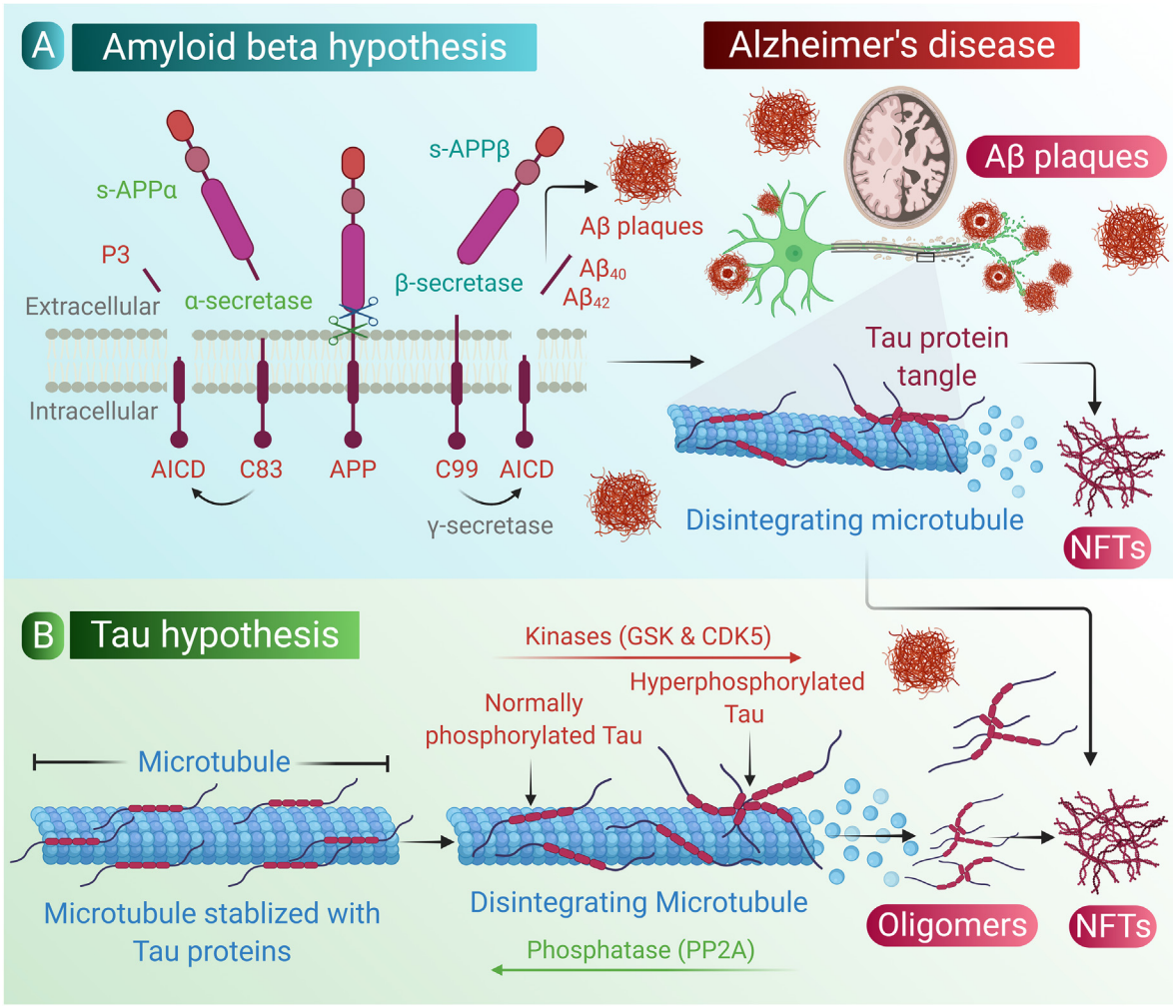


Fig. 1. Schematic showing Amyloid beta and Tau hypothesis that have been suggested to give an explanation for the most common characteristic hallmarks of AD. APP, amyloid precursor protein; AICD, APP intracellular domain; NFTs, neurofibrillary tangles; GSK, glycogen synthase kinase; CDK5, cyclin-dependent kinase 5; PP2A, protein phosphatase 2A.

[14]. Furthermore, a possible molecular link between Aβ deposition and NFT formation has also been reported. The latter has been reported to inhibit neuronal viability, neuroplasticity, and altered microtubule assembly, along with inhibition of mitochondrial transportation along microtubules. Hence, Tau neurotoxicity may be driving an event downstream from Aβ polymerization, and be responsible for the neurotoxicity caused by Tau [15]. However, this hypothesis needs to be experimentally confirmed.

The introduction of the amyloid beta hypothesis provided directions for the development and testing of therapeutic agents for disease modification, which mutually enhance clearance of toxic peptides out of the brain and prevent Aβ formation [16]. Unfortunately, >400 clinical trials conducted over the decade when the last drug for AD was approved have failed. The drug only offered temporary treatment for AD symptoms. A significantly high attrition rate was indicated by an analysis of various clinical trials involving AD between 2002 and 2012, with a mere overall success rate of only 0.4%, with a 99.5% failure rate [11]. Despite the many plausible reasons behind the failure of these clinical trials, a widely-agreed view is that the disease stage was too advanced for any anti-Aβ drugs to impact cognition. Due to the considerable failure on present approaches focused on modifying the disease, other probable treatment methodologies, such as genome editing techniques, are needed. Currently, there are three major prevalent gene editing tools available, including CRISPR/Cas9, transcription activator-like effector nucleases (TALENs), and zinc-finger nucleases (ZFNs). Each of these tools has its own merits and demerits [17]. However, the present review mainly elaborates on the possible role of CRISPR/Cas9 in AD due to its low cost, high speed, efficiency, and precision over other genome editing tools.

CRISPR/Cas9: A promising gene editing tool

CRISPR/Cas9 is a recently discovered and promising revolutionary tool for genome editing, which allows to treat diseases with limited or scarce treatment options. This tool was initially identified by Ishino in 1987 [18] (Fig. 2). Since then, several studies have reported that the CRISPR/Cas9 system is an integral part of a bacterium's immune system, which offers protection from undesired integration of mobile genetic elements such as plasmids and viruses. Further, Doudna and Charpentier's pioneering efforts brought CRISPR/Cas9 to laboratory settings to investigate its potential [19]. In-depth studies on CRISPR/Cas9 have been conducted in the recent past, and significantly improved editing efficiency and minimized off-target effects while being extensively used for basic and translational research [19,20].

CRISPR/Cas9 has two main components: the Cas9 enzyme and single-guide RNA (sgRNA). The target DNA sequence is recognized by the sgRNA, wherein various parameters are considered in the design process to improve specificity. In contrast, as an endonuclease, the Cas9 protein acts as a molecular scissor for incision of the DNA double strands (Fig. 3). CRISPR/Cas systems are classified into

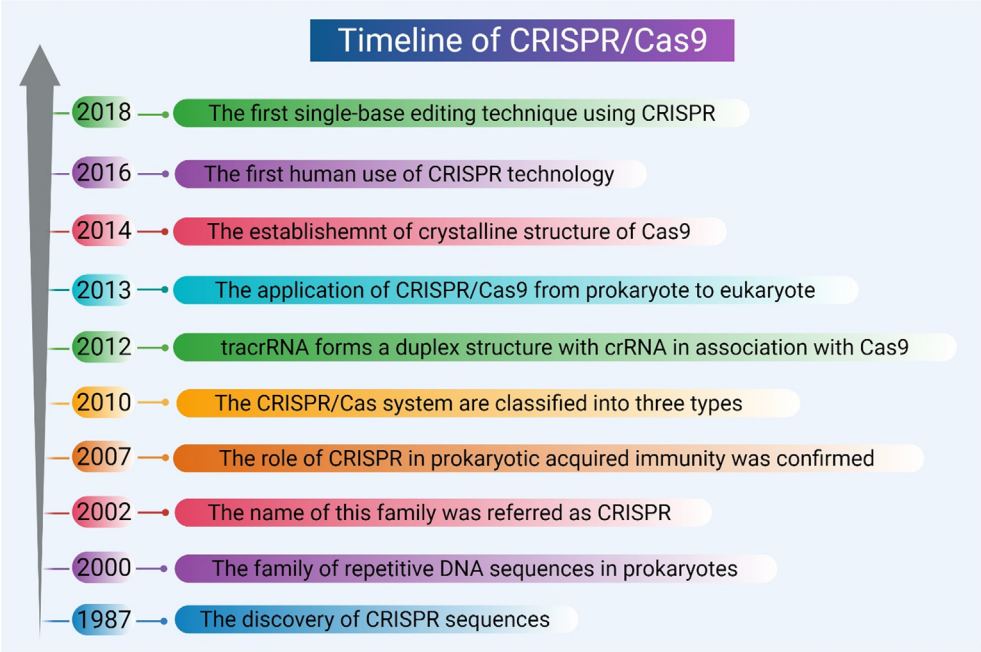


Fig. 2. The CRISPR/Cas9 timeline. crRNA, CRISPR-derived RNA; CRISPR/Cas9, clustered regularly interspaced short palindromic repeats/CRISPR-associated proteins 9 system.

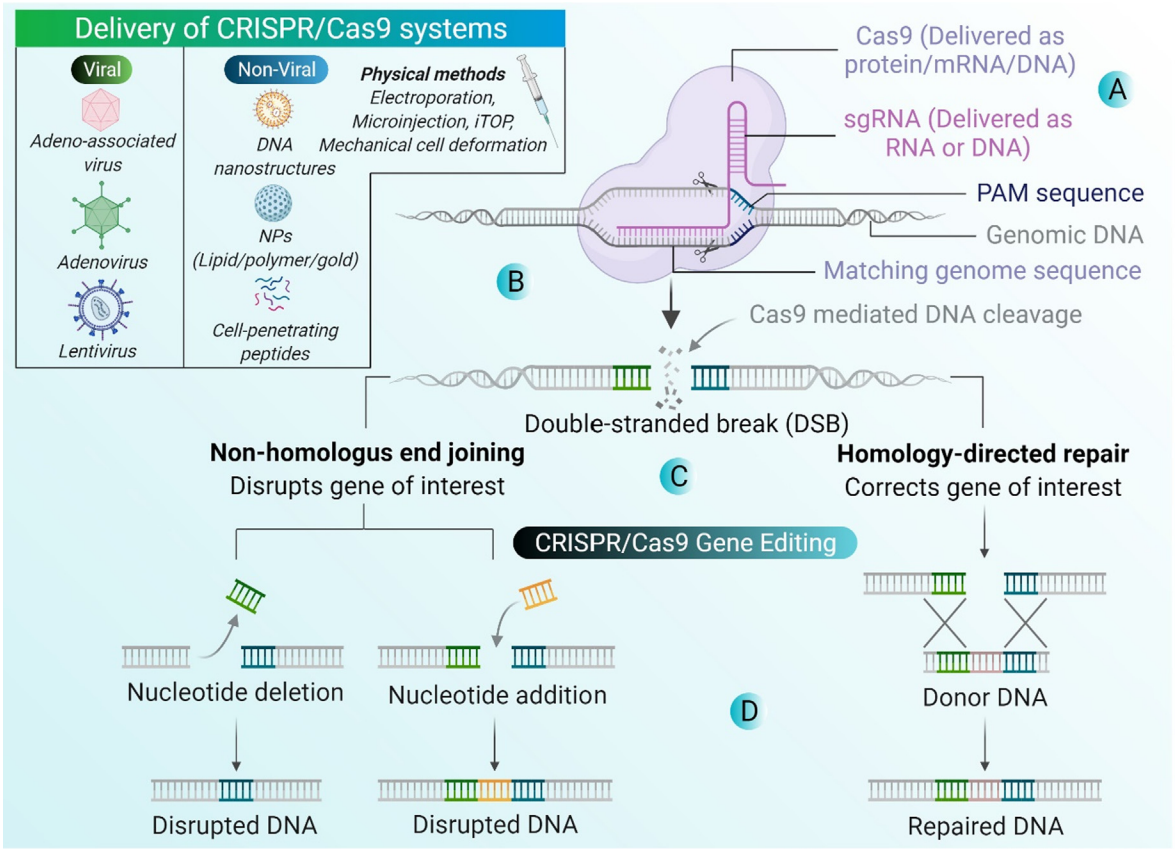


Fig. 3. A schematic cartoon illustrating the steps involved in CRISPR/Cas9 technique. (A) Specially designed sgRNA (guide RNA) which matches with genomic DNA sequence containing mutation, attaches with Cas9 (CRISPR-associated endonuclease), a DNAase capable of inducing a double strand break, thereby, forming Cas9-sgRNA complex. (B) Association of Cas9-sgRNA complex with the target genomic DNA. Cas9 searches for appropriate sequence in target DNA with the help of sgRNA and recognises it with the help of PAMs (protospacer adjacent motifs) sequence, usually 2–6 base pair long, found 3–4 nucleotide downstream from cut site generally serve as a tag. (C) Cas9 mediated DNA cleavage leads to the formation of double strand break (DSB). (D) Formation of DSB leads to the activation of DNA repair mechanism to correct the break by sealing the gap either by Non-Homologous End Joining (NHEJ) or Homology Directed Repair (HDR).

Table 1
Different strategies that have been utilized by CRISPR/Cas9 tool in genome editing.

Different strategies	Principle	Advantages	Disadvantages	Reference(s)
Plasmid-Borne CRISPR/Cas9 System	Plasmid encoding sgRNA and Cas9 protein with suitable promoters	• Fluorescent protein assisted plasmid to label cells expressing Cas9 enzyme • Good reproducibility, economical and extensive adaptability and customizability because sgRNAs can be cloned from same plasmid and may consists of homology directed repair (HDR) template • Improved stability, particularly in handling and manufacturing in comparison with available strategies • Both Cas9 protein and sgRNA are loaded on the same vector; ensures that both are expressed in the same cell	• DNA and bacterial DNA sequences exists in the plasmid elicits cytotoxicity • Potential for the random insertion of plasmid fragments into the gene • Low transfection effciency of primary cells	[23,24]
Cas9/sgRNA Complex	Delivery of sgRNA complexed with Cas9 protein	• Minimal off -target effects and cell toxicity • The complex is fastest among available strategies are not necessary for transcription or translation to Cas9 protein • Ease and convenience due to Cas9 protein spontaneously forms a complex with sgRNA being oppositely charged	• Purification process of the Cas9 protein and free from contamination of endotoxin is expensive • The intracellular Cas9 delivery is complex and challenging due to large size (about 160 kDa)	[25,26]
Cas9 mRNA and sgRNA	Delivery of sgRNA and Cas9 mRNA	• Lower cytotoxicity, off-target effects reported comparatively plasmid-based system • This method is faster over plasmid approach for the editing of the targeted gene because Cas9 mRNA translation is necessary to generate Cas9 protein	• Instability of RNA	[27]

Class 1 (type I, type III, and type IV) and Class 2 (type II, type V, and type VI). Class 1 contains a variety of Cas proteins that work together, whereas Class 2 utilizes a single Cas protein, which makes it simple and desirable for genome editing [21]. Among Class 2, the type II CRISPR/Cas9 system is one of the most researched and utilized systems in pharmaceutical development. The Cas9 protein generates a double standard break after recognizing the target gene sequence. Subsequently, two distinct pathways can be initiated to repair this break, that is, homology-directed repair (HDR) or non-homologous end joining (NHEJ). NHEJ results in insertion and deletion, leading to premature stop codons and/or DNA frameshifts, eventually resulting in gene inactivation, while the HDR pathway helps replacing the mutated/faulty sequence with the correct one. To initiate HDR, thanks to the assistance of a donor DNA template, the right DNA sequences are incorporated into the desired site [22]. In addition, HDR is limited to the G or S phase, while NHEJ can occur in every cell cycle phase. Generally, the HDR pathway provides a highly reliable DNA repair mechanism, though its efficiency is lower than that of the NHEJ pathway. There are three possible ways to edit the intended gene with the help of the CRISPR/Cas9 system: purified Cas9/sgRNA complexes, plasmid-borne CRISPR/Cas9 system, or a combination of Cas9-mRNA and sgRNA. As summarized in Table 1, each strategy has its own merits and disadvantages.

Promising role of CRISPR/Cas9 in AD

Genetic mutations account for approximately 1% of familial cases of AD; thus, genome editing with CRISPR/Cas9 may be useful in familial AD (FAD) at large with minimal or negligible benefits in sporadic AD (SAD). However, considering the involvement of dysregulated Aβ metabolism in FAD and SAD, limiting Aβ production may offer a therapeutic approach independent of the onset, whether familial or sporadic (Fig. 4). Table 2 presents a brief summary of studies that demonstrated the benefits of CRISPR/Cas9 technology as an experimental therapeutic approach in FAD and SAD.

Prospect in early-onset AD models

Treatment with CRISPR/Cas9 might not be appropriate because the onset and progression of most AD cases are sporadic and involve unknown triggers. Actually, only in a few AD cases (<1%), there are real or known mutations in associated genes, that lead to APP production facilitating APP processing to generate Aβ. However, mutations constitute only a small fraction of AD cases, but elicit increased Aβ generation [31]. Similarly, mutations in the presenilin 1 (PSEN1) and presenilin 2 (PSEN2) genes also cause early AD onset [32,33], as they result in accelerated Aβ1-42 production, possibly by altering the APP cleavage site [34]. In most cases, these mutations are usually evident in < 60 years of age and hence categorized as early onset AD. The CRISPR/Cas9 technique can significantly correct these autosomal dominant mutations. Recent studies also support the potential of this gene editing system, which has been reported to correct similar types of mutations. For instance, the CRISPR/Cas9 system was utilized in basal forebrain cholinergic induced pluripotent stem cells (iPSC)-derived neurons from a PSEN2N141I mutated individual for the correction autosomal dominant mutations [35], resulting in correction and stabilization of the Aβ42/40 ratio. Furthermore, the PSEN2 mutation corrected by this editing system also reversed electrophysiological deficits. Previous studies wherein CRISPR/Cas9 was utilized to fix PSEN gene mutations in FAD using iPSCs derived from the patient further supported these results [36,37]. Another study reported that this system helps knocking out Swedish APP mutations in patient-derived fibroblasts finding a 60% Aβ reduction [38]. The Swedish mutation is known to promptly adjoin the β-secretase site in APP [39]. The researchers also interrupted this mutation in Tg2576 mice, which displayed a manifold APP Swedish mutation. To accomplish this, DNA encoding both guide RNAs and Cas9 in AAV vectors was injected into the hippocampus of transgenic mice. Following injections, disruptions such as single-base pair insertions in the APP Swedish gene were observed. However, we need to understand whether such manipulations can improve the behavior and pathology deficits in Tg2576 mice. Notably, when CRISPR/Cas9 was directly injected into the hippocampus, it resulted in only 2% of transgenes disrupted on the injection site [38]. This can be attributed to the fact that Tg2576 mice have approximately 100 transgene copies per neuron, and therefore the injected CRISPR/Cas9 levels appear insufficient to correct the Swedish mutation. While this analysis promises tantalizing developments, an overall more systematic study of the targeted hippocampal cells is required to recognize and improve the editing efficiency and translate the effects *in vivo*.

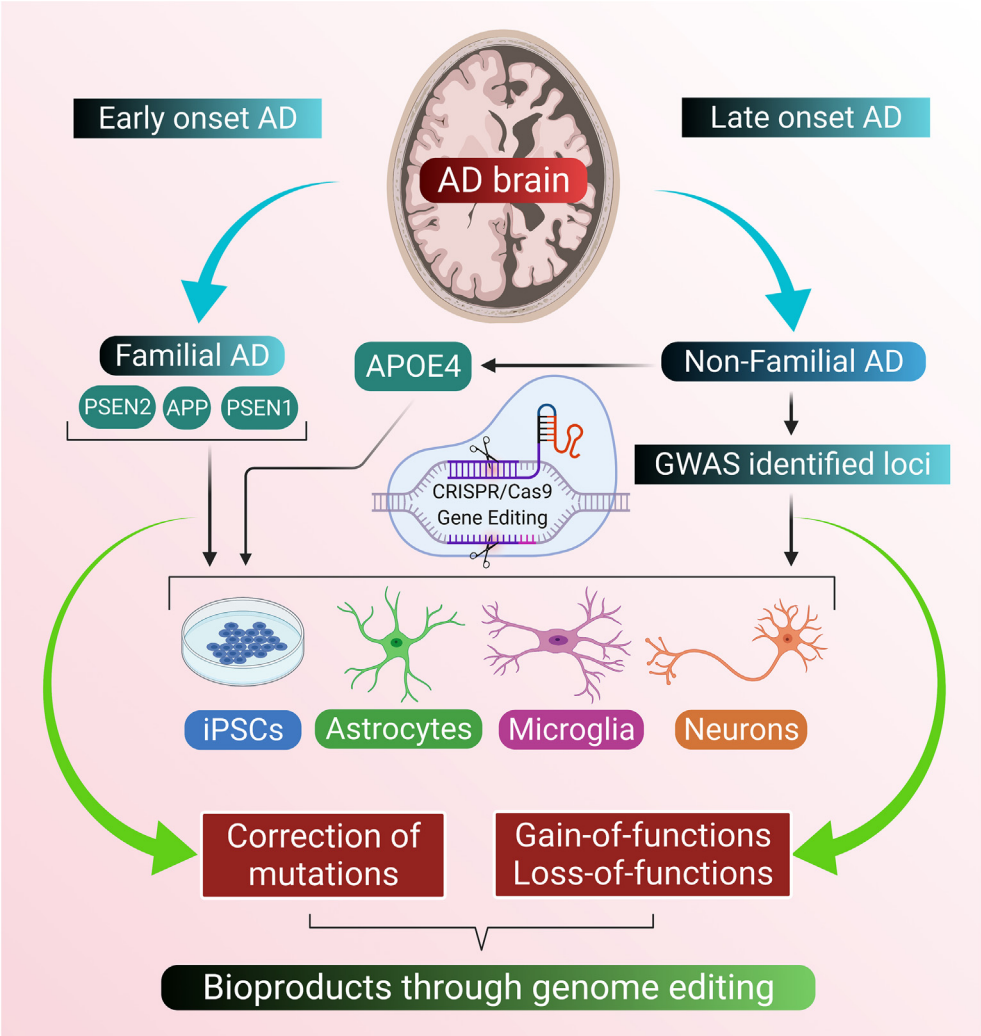


Fig. 4. Illustration showing the possible CRISPR/Cas9 mediated gene editing approach in AD. GWAS, genome-wide association studies.

Table 2
Summary of various reported studies on AD treatment involving CRISPR/Cas9 technique.

Delivery system	Targeted genes	FAD or SAD	Cell lines	Animals tested	Reference(s)
Lentiviral CRISPR/Cas9 system	*APP* at the extreme C-terminus	FAD	-HEK293 -neuro2a cells	Injections administered in the dentate gyrus of *C57BL/6* mice	[28]
R7L10 peptide, a component of nanocomplexes complexed with Cas9-sgRNA ribonucleoprotein	BACE1	FAD	-Mice embryos derived primary neurons -Human induced pluripotent human stem cells and human embryonic stem cells -GFP^{+} HEK293T cells	Injections into the hippocampus:- −*5xFAD* transgenic mice co-expressing 5 FAD mutations - *APP* knock-in transgenic mice	[25]
Recombinant adeno-associated virus (rAAV) mediated delivery of CRISPR/Cas9	*KM670/671NL* APP (APP^{swe}) mutation	FAD	-APP^{swe} fibroblasts in Human -Tg2576 mice embryos derived primary neuronal cells	Injection into hippocampus of *Tg2576* mice	[23]
CRISPR/Cas9 delivered via lentivirus	*APOE E4*	SAD	-Mouse astrocytic cells expressing the human APOE3 or APOE4 gene	NA	[29]
Cytidine deaminase enzyme conjugated CRISPR/Cas9 plasmids	*APOE E4* (converting it into *APOE3r*)	SAD	-HEK293T cells -Immortalized mouse astrocytes containing the APOE E4 isoform of the APOE gene	NA	[30]

Prospective in sporadic AD models

These findings paved the way for evaluation of CRISPR/Cas9 gene editing in subjects with early AD onset, but what about SAD, which by far, represents the vast majority of cases in the USA? To check the efficacy of this gene editing technology in SAD, Sun et al. demonstrated editing of endogenous APP at the extreme C-terminus to mitigate β cleavage and Aβ generation [27]. In doing so, they inhibited further interactions with BACE1 within endosomes, thus avoiding the most important cleavage

event in Aβ generation [40]. Findings in human iPSC-neurons, cultured neurons, cell lines, and mouse brain showed that this limits the physical contact of APP and BACE1 and therefore mitigates Aβ generation [28]. The APOE4 allele is another significant risk factor contributing to late-onset AD [41]. APOE exists in three isoforms: APOE2, APOE3, and APOE4. All of them vary only by substitution of a single amino acid leading to the replacement of cysteine-arginine at positions 112 and 158 [42]. And different isoforms have different properties, for instance, the possibility of developing AD is reduced by up to 40% by one copy of the E2 allele, which is the rarest form of APOE. APOE3, one of the common alleles, does not appear to affect the risk of AD, while APOE4 present in approximately 10–15% of the people, reduces the age of AD onset, and enhances the risk of AD [43]. Risk, which can increase by 2–3-fold if individuals carry a single copy of E4 (E3/E4), whereas double copies of E4 (E4/E4) may enhance it by approximately 10–15 times. Actually, 65–80% of all patients diagnosed with AD have $\geq$ 1 APOE4 allele [43,44]. Although several APOE4 adverse effects appear linked with Aβ accumulation, a recent study suggests that Tau phosphorylation in neurons derived from human iPSCs may be promoted by APOE4, independent of Aβ [45,46]. In this study, gene editing by zinc-finger nuclease to converts APOE4 to APOE3, prevented APOE4-linked pathology in their model system [46]. Hence, CRISPR/Cas9 could also act as a potential editing tool to transform APOE4 to APOE2 or E3. Notably, a study points towards important APOE4 structural features, which differentiate it from APOE2 to E3, and involves a domain interaction facilitated through a salt bridge between amino acids Arg-61 and Glu-255 [47]. Thus, altering one of these amino acids with CRISPR/Cas9 can effectively neutralize the risk related to the APOE4 allele. Further, Table 3 addresses the therapeutic role of CRISPR/Cas9 in correcting specific gene sequences.

CRISPR/Cas9 delivery system: A possible way to target AD

In the management of AD, CRISPR/Cas9 genome-editing holds promise. Although efficient and safe delivery systems are still lacking, it is a vast task that requires translation of this technological approach into real therapeutic applications. To date, viral and non-viral methodologies for the delivery of the CRISPR/Cas9 systems are available.

Viral vectors for CRISPR/Cas9

The use of viral vectors for delivering CRISPR/Cas9 is a classical method in experimental models, including cell lines and animals, and one of the most dynamic systems for the targeted delivery of plasmid-based CRISPR/Cas9. Moreover, they may incorporate mutations that have significant adverse effects. Adeno-associated virus (AAV) are a frequently used vector because of its high infectivity, low immunogenicity, and low integration into the human genome [24,25]. With > 200 variants, the AAV genome comprises single-stranded DNA [26]. One study reported the use of two separate AAV vectors packaging APPsw-specific gRNA and Cas9 targeting the AD-causing KM670/671NL APP mutation. The viruses were tested *in vitro* in primary neuronal cells from Tg2576 mouse embryos and *in vivo* via intrahippocampal injection in Tg2576 mice. This treatment reduced Aβ generation by approximately 60% in human fibroblasts [38]. Due to the low packing capability of the 4.7 kb of AAV, co-injections of two viruses might be required. However, this would further complicate the procedure as both can not concurrently infect the same cell. Lentivirus incorporate long DNA inserts comprising 8–10 kb but with lower brain disseminating efficiency [27]. But, contrary to AAV, lentiviruses cannot be so easily produced in large quantities, and can be incorporated into the human genome, thereby provoking immune reactions [25]. However, studies have demonstrated the possible use of lentivirus to target three genes, namely, APOE4, APP, and caspase-6, in SAD and familial AD [28-30].

Non-viral vectors for CRISPR/Cas9

Non-viral vectors have been found promising for the targeted delivery of CRISPR/Cas9, attributed to better cost effectiveness, relative ease, feasibility, and flexibility. Consequently, they are extremely suitable for application in AD. It is very easy to form nanocomplexes by combining positively charged CRISPR/Cas9 peptides with negatively charged nucleic acid cargo. Comparatively, they are less immunogenic than viral vectors; they can also assist in numerous applications, as they are compatible with ligands. Although delivering nanocomplexes into the brain is challenging, as they are unable to properly cross the blood brain barrier (BBB) via systemic administration, and the reticuloendothelial system (RES) also actively removes them from the blood. Hence, intracerebroventricular and intrathecal injections were used as standard. However, in direct injection methods, multiple injections are needed to ensure proper distribution across the brain, thereby restricting their application. Park et al. used a Cas9–sgRNA ribonucleoprotein specifically targeting BACE1 complexed with nanocomplexes made of R7L10 peptide [49], and reported successful targeting, reducing BACE1 expression without any substantial off-target mutation *in vivo*. In addition to CRISPR/Cas9 delivery, several other vehicles can transport short interfering RNAs (siRNAs) for targeting AD across the BBB. Recent reports have used polymeric nanocomplexes of Poly(mannitol-polyethyleneimine) (PMT) carrier amended with rabies infection glycoprotein (RVG) [27]. The polymer was designed to form a complex with siRNA to target the BACE1 gene. The nanocomplexes were intended to possess enhanced transmission ability owing to the presence of the RVG ligand, which ameliorates permeation across the BBB and focuses on nerve cells. Polyethylene glycol (PEG) reduces transfection efficiency by creating a positively charged protection that prevents linking to cell membranes. It was proposed that this issue could be resolved using the RVG ligand, thereby ameliorating the cellular uptake of the nanocomplexes. The decrease in Aβ1-42 cortical levels further validated the silencing ability of nanocomplexes. However, considerable therapeutic potential loss may result from unknown body distribution, therefore, the suitability of such a delivery approach should be investigated in different AD models.

The following systems are promising for applications in AD. DNA nanoclews could deliver the Cas9–sgRNA complex. The traditional assembly of DNA nanostructures based on base-pairing is complex and time-consuming. Sun et al. reported that DNA nanoclews are confined nanosized DNA moieties that comprise polyethylenimine for applying a positive charge for endosomal seepage and improved cellular uptake [60]. Nanoclews bearing the sgRNA-Cas9 complex targeting enhanced green fluorescent protein (EGFP) were injected locally into tumor-bearing mice. After ten days of treatment, this resulted in a 25% reduction in EGFP expression [60]. However, nanoclews might induce an immunogenic response, which needs to be studied further. Importantly, polymeric nanoparticles and lipid nanoparticles also exhibit potential as CRISPR/Cas9 delivery tools. These nanoparticles have been largely utilized for carrying gene editing tools in hepatitis, cancer, and several viral infections [61,62]. However, their potential use in AD requires further investigation. In addition, Wang et al. used AuNPs for investigation [63]. The CXCR4 gene targeted with CRISPR-Gold achieved HDR efficiency about 3–4% in many types of human cells. Furthermore, local infusion of CRISPR Gold into

the gastrocnemius and tibialis frontal muscles of mdx mice resulted in correction of the mutated dystrophin gene, causing inborn Duchenne muscular dystrophy [63]. However, after CRISPR Gold injection, no meaningful changes in the inflammatory cytokine profile were found, indicating low toxicity. Recently, attention has been drawn to the therapeutic delivery of CRISPR/Cas9 through microvesicles. Typically, a "producer" cell line is transfected with a microvesicle-prompting protein (RAB proteins), sgRNA, and Cas9 proteins [64]. The cells produce microvesicles consisting of Cas9–sgRNA complexes, which are shed into a medium subsequently sterilized and used to deliver gene editing tools to the target cells.

CRISPR/Cas9 technique: A boon in the bag for neurodegenerative diseases

In the previous section, we extensively elaborated on the possible role of CRISPR/Cas9 in the treatment of AD. However, it is also important to address the implications of this technology in other NDDs and for the management and treatment of brain abnormalities. Recently, this technology has been implemented in the treatment of several neurodegenerative disorders, including PD and HD, as discussed in the following sections.

Parkinson's disease (PD)

PD is an age-related, progressive, multifactorial and one of the most common neurodegenerative condition, manifested by motor and non-motor motor indications. Proper treatment for PD is still lacking because the pathological mechanism and signaling axis responsible for disease progression are not fully understood. However, CRISPR/Cas9 might be helpful in the identification of pathways and proteins associated with PD pathogenesis (Table 4). Additionally, this technology may also help in the identification of complex interactions between human genetics and environmental factors leading to PD [65]. Single nucleotide polymorphisms (SNP) variants in the SCNA gene (α-synuclein) were analyzed by Soldner et al. using this system, and they reported the presence of a non-coding distal enhancer element common variant responsible for the enhanced SCNA expression [66]. This tool also helps scientists in PD research to produce isogenic cell lines for PD modeling and hence could contribute to analyzing PD phenotypes. In this regard, Arias-Fuenzalida et al. applied fluorescent markers and the CRISPR/Cas9 system to obtain biallelic genome-edited cell populations. They called this approach FACS-assisted CRISPR/Cas9 editing (FACE). The FACE method helps in deriving a set of isogenic cell lines with mutated α-synuclein associated with PD [67]. On the other hand, mutations in leucine-rich repeat kinase 2 (LRRK2) are the most common genetic cause of sporadic and familial PD, which causes toxicity in dopaminergic neurons. The CRISPR/Cas9 tool was used to edit mutated LRRK2 and reduced neurite complexity in dopaminergic neurons and the incidence of both sporadic and familial PD [68].

CRISPR has also recently been recognized as a promising system to assist in the understanding of the interplay between PD genes and the identification of new apoptotic cascades directly or indirectly associated with PD. For instance, a Parkin regulator, THAP11, was recently validated in CRISPR/Cas9 knockout research in various cell types, to find previously undiscovered regulatory crosstalk/networks [69]. This system has also been employed to study the neuroinflammatory mechanisms associated with PD. For instance, protein kinase Cδ (PKCδ) signaling linked with Mn-induced apoptotic cell death in PD involves PKCδ activation. CRISPR/Cas9-mediated PKCδ downregulation in dopaminergic neurons (DA) neurons, considerably hindered the DNA fragmentation induced by Mn. In addition, this system can also help understand the PKCδ downstream pathway driving apoptosis [70].

The role of CRISPR/Cas9 was also reported by Gordon et al. in the elimination of Prokineticin-2 (PK2). In PD, altered PK2 has been observed to enhance neuronal susceptibility to neurotoxic-stimulated cell death following neuroinflammatory mechanisms. They also reported higher expression of PK2 in PD post-mortem brains, and served as a protective compensatory response against neurodegeneration in cell culture and PD animal models [71].

Glia maturation factor (GMF), a neuroinflammatory and microglia-associated protein is abundantly expressed in the brain, with increased expression in the substantia nigra (SN) of PD brains [72]. Selvakumar et al. applied this system to investigate the impact of GMF editing in microglial cells under oxidative stress conditions and Nrf2/HO-1-dependent ferritin activation. They noticed that GMF knockout in microglial cells weakened oxidative stress by curtailing reactive oxygen species (ROS) production and lowering calcium flux. Moreover, the absence of GMF reduced Nrf2 nuclear translocation, thus preventing microglia activation

Table 3
Overview of reported clinical trials on AD therapeutics using CRISPR/Cas9 technique.

Mutations may fixed with CRISPR/Cas9	Targeted genes for CRISPR/Cas9	Clinical outcomes	Reference(s)
PSEN2^{N141I} mutation	*PSEN2*	Decreased Aβ42/40 ratio	[48]
Manipulation in Aβ-linked pathologies	*BACE1*	Considerable downregulation of Aβ42 plaque aggregation in mice	[49]
NHEJ-mediated exon removal	*MAPT*	Production of new Tau knockout strain (tauΔex1) in mice	[50]
HDR-mediated mutation	*PSEN1M1*	Disease models generated by CRISPR	[51]
HDR-mediated mutation	*APPS*	Disease models generated by CRISPR	[51]
Reciprocal manipulation of the amyloid pathway	*APP*	Attenuation of β-cleavage and Aβ production	[28]
N141I	*PSEN2*	Normalization of enhanced levels of Aβ42/40 via CRISPR/Cas to correct mutation in PSEN2^{N141I}	[35]
Met146Val	*PSEN1*	Establishment of homozygous and heterozygous mutations	[52]
Conversion of APOE E4 to APOE E3	*APOE*	Arg158 converted to Cys158 in 58–75%	[53]
Mutations in L52P, T48P and K53N	*APP*	A model to investigate outcomes of APP mutations in cleavage of γ-secretase and Notch signaling	[54]
Deletion of Swedish mutation	*APP*	Reduction in *ex vivo* and *in vivo* production of Aβ peptide	[38]
Glia maturation factor (GMF)	*GMF*	Reduction in GMF and p38 MAPK	[55]
3′-UTR amyloid precursor protein (APP)	*APP*	Reduction in APP and Aβ	[56]
γ-Secretase activating protein (GSAP)	*GSAP*	Reduction in GSAP, γ-Secretase activity and Aβ	[57]
β-secretase 1 (BACE1) and Tyrosine hydroxylase (Th)	*BACE and Th*	Reduction in BACE1, Th1 and Aβ	[58]
APOE-E3/E4 APOE	*APOE*	Reduction in APOE-E3/E4 (high risk for AD), hyper-phosphorylation of Tau protein and amyloid deposition and upregulation in APOE-E2/E2 (low risk for AD) and turning APOE4 to APOE3 is increased	[59]

Table 4
Overview of reported clinical trials on PD and HD therapeutics using CRISPR/Cas9 technique.

Target genes	Mechanisms	Main findings	*In vitro* or *in vivo*	Disease	Reference (s)
SNCA	CRISPR-mediated SNCA deletion in human embryonic stem cells (hESCs)	Reduction in SNCA alleles	hESCs		[82]
-Parkin -SNCA	CRISPR/Cas9-based gene targeting in α-Syn preformed fibrils (PFF cells)	The use of gene-targeted somatic cells as a donor for somatic cell nuclear transfer (SCNT) to generate gene-targeted animals with single and identical mutations	PFF cells		[83]
SNCA	Fluorescent markers derived biallelic genome-editing	Examine a set of isogenic lines comprising PD-linked α-Syn in mutation	*In vitro*	PD	[84]
LRRK2	The Cre-LoxP recombination system has been utilized to alter the LRRK2-G2019S mutation in the human induced pluripotent stem cells (hiPSCs) to produce isogenetic controls	Production of an edited footprint-free LRRK2-G2019S isogenic hiPSCs	hiPSCs		[68]
-DJ-1 -Parkin -PINK1	Co-injection of multiplexing sgRNAs and Cas9 mRNA into *in vivo* derived pronuclear embryos	Reduction in DJ-1, Parkin, and PINK1	*In vivo*		[85]
The polyglutamine repeat in the huntingtin gene (HTT)	CRISPR/SpCas9 technique	Non-allele specific CRISPR/Cas9 mediated permanent elimination of polyQ domain of mHTT to reduce neuronal toxicity in the adult brain	Mouse striatum		[86]
Single nucleotide polymorphism (SNP) sites	Allele specific CRISPR/Cas9	Specific CRISPR/Cas9 alleles mediated permanent inactivation of Huntington's mutation allele	Fibroblasts cells	HD	[87]
Transient receptor potential canonical 1 (TRPC1)	CRISPR/Cas9 system	CRISPR/Cas9 mediated TRPC1 Inhibition may serve as a neuroprotective tactic in the treatment of disease	YAC128 HD mice/ HEK293T cell line		[88]
The SNPs at upstream and downstream (Intron1) of HTT exon-1	CRISPR/SpCas9 technique	SNPs are identified that either accountable for causing or destroying PAM motifs critical for CRISPR-selective editing of one allele versus the other in cells from HD patients and in a transgenic HD model harboring human allele	Fibroblast and HEK293 cell		[77]
5′-DNA at the uORF region/exon1-intron region of the mHTT	CRISPR/Cas9 mediated silencing of the mHTT gene *in vitro*	Reduction in mHTT production	Plasmids used in MSCs extracted from the YAC128 mice bone marrow, which carries the transgene for HD		[89]

and reducing the expression of pro-inflammatory molecules in the brain [73].

Both monogenic recessive and dominant-negative disorder-induced mutations could be corrected or inactivated by the CRISPR/Cas9 system. In DA neurons of PD animal models, CRISPR/Cas9 has been reported to inactivate and deplete mutated genes that express mutant SNCA [74]. For instance, Chen et al. used this system to delete the SNCA gene in hESCs and slowed the progress of this pathological outcome associated with SNCA [75]. However, more clinical research is still required to validate these results and find the proper treatment of PD involving this technology.

Huntington's disease (HD)

HD is an autosomal dominant trinucleotide repeat neurodegenerative disorder caused by the expansion of CAG repeats in the Huntington (HTT) gene, which encodes a mutant HTT protein (mHTT). Expansion of the CAG trinucleotide segment consists of glutamine residue repeats (PolyQ), which makes the mutant HTT protein longer than normal. CRISPR/Cas9 can selectively suppress mHTT expression through direct interaction with the DNA. Furthermore, the production of mHTT can be reduced by suppressing its endogenous expression in the striatum of mHTT-expressing mice [76]. In another study, SNPs with either causative or destructive effects on PAM motifs were critical in the selection of one allele for CRISPR editing vs. the other to enhance the efficiency of Cas9 nuclease and apply CRISPR strategy for therapeutic purposes of HD [77]. The presence of the mHTT gene leads to sensitization of type 1 inisitol 1,4,5-triphosphate receptor (InsP3R1) and, consequently, calcium outflow from the ER and a compensatory elevation in neuronal store-operated calcium (nSOC) entry, ultimately resulting in synaptic damage of striatal MSNs in an HD animal model [78]. A transient receptor potential canonical 1 (TRPC1) knockout (which is one of the nSOC components) using CRISPR/Cas9 resulted in improved motor performance and salvage of MSN spines *in vivo* and *in vitro* [79]. Further studies proposed for the regulation of HTT protein synthesis, the 5 untranslated region (UTR) plays a major role as it contains an uORF that encodes 12 amino acid long potential polypeptide that controls downstream ORF translation. Hence, the expression of Huntingtin mRNA is negatively affected by the presence of uORFs [80]. Likewise, previous studies have proposed that disruption of the uORF in the 5UTR of mRNA using CRISPR/Cas9 may lead to decreased translational products of mutant huntingtin gene in MSCs derived from HD mouse models. Kolli et al. studied mHTT inhibition with two types

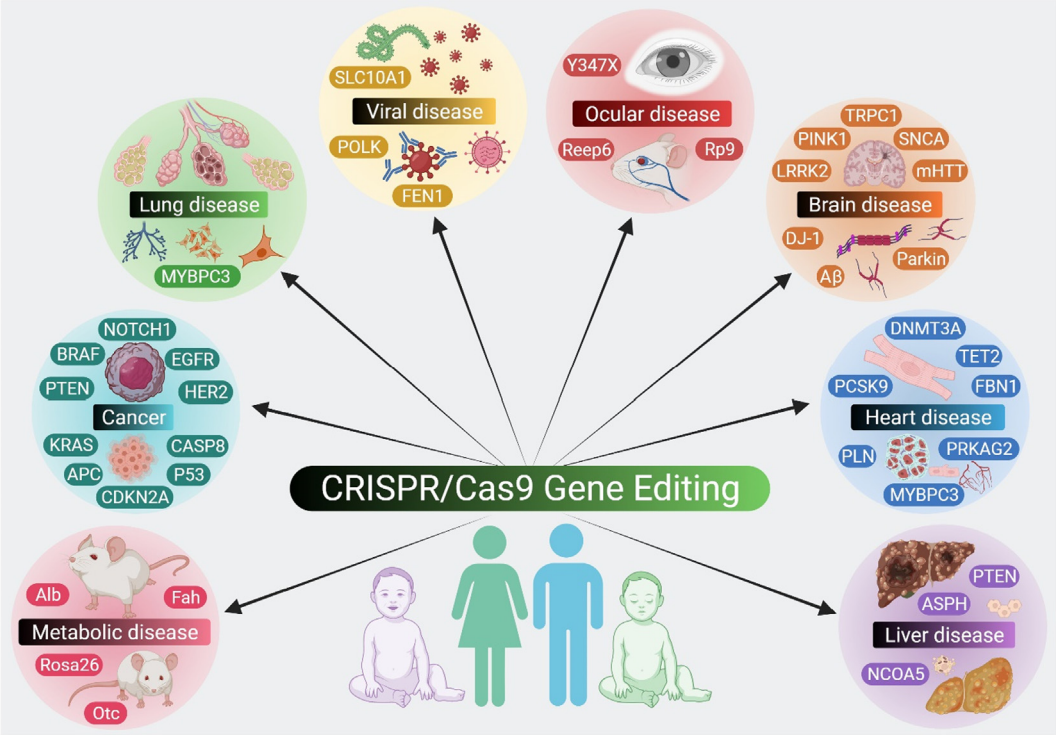


Fig. 5. The diverse applications of CRISPR/Cas9 technique in human diseases. Here in this schematic, we have highlighted the disease associated various genes and proteins which may be one of the possible target for this gene editing strategy. MYBPC3, myosin binding protein C3; BRAF, B-Raf proto-oncogene, serine/threonine kinase; PTEN, phosphatase and tensin homolog; EGFR, epidermal growth factor receptor; HER2, human epidermal growth factor receptor 2; CASP8, caspase-8; CDKN2A, cyclin-dependent kinase inhibitor 2A; SLC10A1, solute carrier family 10 member 1; TRPC1, canonical transient receptor potential; PINK1, PTEN-induced putative kinase 1; LRRK2, leucine rich repeat kinase 2; SNCA, α-synuclein; mHTT, mutant huntingtin protein; DJ-1, PARK7; DNMT3A, DNA methyltransferase 3a; FBN1, fibrillin-1; PCSK9, proprotein convertase subtilisin/kexin type 9; PLN, phospholamban; PRKAG2, kinase AMP-activated noncatalytic subunit-2; ASPH, aspartate beta-hydroxylase; KRAS, kirsten rat sarcoma viral oncogene homolog; APC, adenomatous polyposis coli; p53, tumor suppressor gene; FEN1, flap endonuclease 1; TET2, epigenetic modifier enzyme; NCOA5, nuclear receptor coactivator 5; Y347X, nonsense point mutation; Reep6, receptor expression-enhancing protein 6; Rp9, pre-mRNA splicing factor; Alb, albumin; Fah, fumarylacetoacetate hydrolase; Otc, Ornithine transcarbamylase; POLK, DNA Polymerase Kappa.

of CRISPR/Cas9 system, one that cuts the DNA at exon1 – intron boundary, and the other in the untranslated uORF region, and found that CRISPR/Cas9 mediated mHTT silencing affects the mHTT translational process and reduces mHTT production in bone-marrow-derived mesenchymal stromal cells (BM-MSCs) [81]. Further, Table 4 addresses the therapeutic potential of CRISPR/Cas9 in NDDs such as PD and HD.

Concluding remarks, major challenges and future perspectives

As a gene-editing tool, CRISPR/Cas9 holds promise by correcting specific gene sequences and having significant potential for the treatment of AD and other human diseases [90-96] (Fig. 5). In SAD and FAD, an altered Aβ metabolism is commonly observed, irrespective of genetic factors. Thus, CRISPR/Cas9 technology could correct increased Aβ production or mutations in APP, PSEN-1, and PSEN-2, as mutations in these genes are a causative factor of FAD.

There are numerous challenges related to the efficiency of AD management following CRISPR/Cas9 brain delivery via non-viral vectors. Preferably, the vectors should be steady and able to efficiently carry the load to the desired site. When the targeted cells are approached by the vectors, they should be internalized to avoid lysosomal degradation and target the nucleus.

It is essential to consider the large size of CRISPR/Cas9 for future designs and application. Instead of plasmid-assisted delivery approaches, the Cas9-sgRNA complex is preferred because of its smaller size. Moreover, owing to the circulating proteases and nucleases, the constituents of the developed formulations are vulnerable to degradation. Although widely employed to lessen the recognition of these systems by RES, PEGylation can generate specific PEG-antibodies and reduce cellular uptake, leading to immunogenic responses [97]. For *in vivo* applications, non-viral vectors are preferred, but improving the diverse formulation features limits real-life applications.

The systemic route is widely studied because of its *in vivo* feasibility, specifically in the case of AD patients, depite the issues regarding delivery vector's stability and targetability. Hence, intracerebroventricular and intrathecal injections are usually adopted for administration. Stereotaxic microinjection surgery has been used for gene delivery in the brains of experimental animals with PD [98]. However, for AD, this procedure can be challenging because of the widespread nature of Aβ pathology. Another promising approach is the intranasal route as it bypasses the BBB. However, more clinical studies are needed on the nasal delivery of CRISPR/Cas9 based therapeutics.

To ensure the safety of CRISPR/Cas9 therapeutics, further research is essential, as genome editing is irreversible. Furthermore, studies investigating possible off-targets and long-term effects are still insufficient, and ethical considerations are needed before any application in humans. Importantly, the CRISPR/Cas9 tool affects somatic rather than germline cells. Hence, gene editing would not be transferred to subsequent generations and only manifest in the individuals undergoing treatment [23].

Although CRISPR/Cas9 involves double-stranded DNA breaks, the recent prime editing can rectify gene mutations without double-strand breaks. Instead, prime editing involves an impaired Cas9 following catalytic processes bonded to a reverse transcriptase, with guidance provided through a prime-editing

guide RNA (pegRNA). Thus, directing the system to the DNA site targeted for the required correction [64]. Further studies are required to establish off-target and potential of this new technology.

Compliance with Ethics Requirements

This article does not contain any studies with human or animal subjects.

CRediT authorship contribution statement

Shanu Bhardwaj: Writing - review & editing, Formal analysis. **Kavindra Kumar Kesari:** Writing - review & editing, Formal analysis. **Mahesh Rachamalla:** Writing - review & editing, Formal analysis. **Shalini Mani:** Writing - review & editing, Formal analysis. **Ghulam Md. Ashraf:** Writing - review & editing, Formal analysis. **Saurabh Kumar Jha:** Writing - review & editing, Formal analysis. **Pravir Kumar:** Writing - review & editing, Formal analysis. **Rashmi K. Ambasta:** Writing - review & editing, Formal analysis. **Harish Dureja:** Writing - review & editing, Formal analysis. **Hari Prasad Devkota:** Writing - review & editing, Formal analysis. **Gaurav Gupta:** Writing - review & editing, Formal analysis. **Dinesh Kumar Chellappan:** Writing - review & editing, Formal analysis. **Sachin Kumar Singh:** Writing - review & editing, Formal analysis. **Kamal Dua:** Writing - review & editing, Formal analysis. **Janne Ruokolainen:** Writing - review & editing, Formal analysis. **Mohammad Amjad Kamal:** Writing - review & editing, Formal analysis. **Shreesh Ojha:** Conceptualization, Writing - original draft, Writing - review & editing, Visualization, Formal analysis. **Niraj Kumar Jha:** Conceptualization, Writing - original draft, Writing - review & editing, Visualization, Formal analysis.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We are grateful to the United Arab Emirates University, UAE for the research grant support and also express gratitude to the unknown referees for their time and valuable suggestions to improve the manuscript.

Authors contributions

NKJ and SO conceptualized the study and hypotheses. SB, KKK, MR, and SM performed literature search. NKJ draw the schemes and drafted the artwork. DKC, SKS, KD, JR, MAK drafted the tables. NKJ, and other authors contributed significantly in editing the manuscript. GMA, SKJ, PK, RKA, HD, HPD and GG significantly contributed during revision. All authors read, edited and approved the manuscript.

Funding

Authors received no specific funding for this work. The author (SO) is grateful to the United Arab Emirates University, UAE for research grant supports.

Consent for publication

All authors have read the final version of the manuscript and have given their consent for publication.

References

[1] Kumar P, Jha NK, Jha SK, Ramani K, Ambasta RK. Tau phosphorylation, molecular chaperones, and ubiquitin E3 ligase: clinical relevance in Alzheimer's disease. J Alzheimers Dis 2015;43(2):341–61. doi: https://doi.org/10.3233/JAD-140933. PMID: 25096626.
[2] Lanctôt KL, Amatniek J, Ancoli-Israel S, Arnold SE, Ballard C, Cohen-Mansfield J, et al. Neuropsychiatric signs and symptoms of Alzheimer's disease: New treatment paradigms. Alzheimers Dement (N Y) 2017 Aug 5;3(3):440–9. doi: https://doi.org/10.1016/j.trci.2017.07.001. PMID: 29067350; PMCID: PMC5651439.
[3] Pandey G, Ramakrishnan V. Invasive and non-invasive therapies for Alzheimer's disease and other amyloidosis. Biophys Rev. 2020 Oct;12(5):1175-1186. doi: 10.1007/s12551-020-00752-y. Epub 2020 Sep 15. PMID: 32930962; PMCID: PMC7575678.
[4] Jha NK, Sharma A, Jha SK, Ojha S, Chellappan DK, Gupta G, et al. Alzheimer's disease-like perturbations in HIV-mediated neuronal dysfunctions: understanding mechanisms and developing therapeutic strategies. Open Biol 2020;10(12):200286.
[5] Jha NK, Jha SK, Kar R, Nand P, Swati K, Goswami VK. Nuclear factor-kappa β as a therapeutic target for Alzheimer's disease. J Neurochem 2019;150(2):113–37.
[6] Jha NK, Jha SK, Kumar D, Kejriwal N, Sharma R, Ambasta RK, et al. Impact of Insulin Degrading Enzyme and Neprilysin in Alzheimer's Disease Biology: Characterization of Putative Cognates for Therapeutic Applications. J Alzheimers Dis 2015;48(4):891–917.
[7] Jha NK, Kar R, Niranjan R. ABC Transporters in Neurological Disorders: An Important Gateway for Botanical Compounds Mediated Neuro-Therapeutics. Curr Top Med Chem 2019;19(10):795–811.
[8] Jha NK, Jha SK, Sharma R, Kumar D, Ambasta RK, Kumar P. Hypoxia-Induced Signaling Activation in Neurodegenerative Diseases: Targets for New Therapeutic Strategies. J Alzheimers Dis 2018;62(1):15–38.
[9] Jha SK, Jha NK, Kumar D, Sharma R, Shrivastava A, Ambasta RK, et al. Stress-Induced Synaptic Dysfunction and Neurotransmitter Release in Alzheimer's Disease: Can Neurotransmitters and Neuromodulators be Potential Therapeutic Targets? J Alzheimers Dis 2017;57(4):1017–39.
[10] Gaugler J, James B, Johnson T, Marin A, Weuve J. Alzheimer's disease facts and figures. Alzheimers Dement 2019;15:321–87.
[11] Cummings JL, Morstorf T, Zhong K. Alzheimer's disease drug-development pipeline: Few candidates, frequent failures. Alzheimers Res Ther 2014;6:37.
[12] Yan R, Vassar R. Targeting the secretase BACE1 for Alzheimer's disease therapy. Lancet Neurol 2014;13:319–29.
[13] ŠimićG,BabićLeko M, Wray S, Harrington C,DelallelJ,Jovanov-MiloševićN, Bažadona, D,Buée L, De Silva R, Di Giovanni G, et al. Tau protein hyperphosphorylation and aggregation in Alzheimer's disease and other tauopathies, and possible neuroprotective strategies. Biomolecules 2016;6:6.
[14] Castillo-Carranza DL, Nilson AN, SkikeCE Van, Jahrling JB, Patel K, Garach P, et al. Cerebral microvascular accumulation of tau oligomers in Alzheimer's disease and related tauopathies. Aging Dis 2017;8:257..
[15] Brandt R, Bakota L. Microtubule dynamics and the neurodegenerative triad of Alzheimer's disease: The hidden connection. J Neurochem 2017;143:409–17.
[16] Galimberti D, Scarpini E. Disease-modifying treatments for Alzheimer's disease. Ther Adv NeurolDisord 2011;4:203–16.
[17] Gupta RM, Musunuru K. Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. J Clin Invest 2014;124:4154–61.
[18] IshinoY Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. J Bacteriol 1987;169:5429–33.
[19] JinekM Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. Science 2012;337:816–21.
[20] Ran FA, Hsu PD, Lin CY, GootenbergJS Konermann S, DA TrevinoAE Scott, Inoue A, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell 2013;154:1380–9.
[21] Mir A, Edraki A, LeeJ Sontheimer EJ. Type II-C CRISPR-Cas9 biology, mechanism, and application. ACS Chem Biol 2018;13:357–65.
[22] ZaboikinM ZaboikinaT, FreterC Srinivasakumar N. Non-homologous end joining and homology directed DNA repair frequency of double-stranded breaks introduced by genome editing reagents. PLoS ONE 2017;12:e0169931.
[23] TozzoP,ZulloS,CaenazzoL. Science Runs and the Debate Brakes: Somatic Gene-Editing as a New Tool for Gender-Specific Medicine in Alzheimer's Disease. Brain Sci. 2020;10:421.
[24] Gaj T, Epstein BE, Schaffer DV. Genome engineering using adeno-associated virus: basic and clinical research applications. Mol Ther 2016;24(3):458–64.
[25] Recchia A, Perani L, Sartori D, Olgiati C, Mavilio F. Site-specific integration of functional transgenes into the human genome by adeno/AAV hybrid vectors. Mol Ther 2004;10(4):660–70.
[26] Grimm D, Kay MA. From virus evolution to vector revolution: use of naturally occurring serotypes of adeno-associated virus (AAV) as novel vectors for human gene therapy. Curr Gene Ther 2003;3(4):281–304.

[27] Dissen GA, McBride J, Lomniczi A, et al. Using lentiviral vectors as delivery vehicles for gene therapy. Controlled genetic manipulations. Berlin: Springer; 2012. p. 69–96.
[28] Sun J, Carlson-Stevermer J, Das U, et al. CRISPR/Cas9 editing of APP C-terminus attenuates b-cleavage and promotes a-cleavage. Nat Commun. 2019;10(1):1–1.
[29] Offen D, Angel A, Ben-Zur T. Caspase-6 knock-out using CRISPR/Cas9 improves cognitive behavior in the 3xTg mouse model of Alzheimer's disease. Cytotherapy 2018;20(5):S94–S9595.
[30] Offen D, Rabinowitz R, Michaelson D, Ben-Zur T. Towards gene-editing treatment for Alzheimer's disease: ApoE4 allele-specific knockout using a CRISPR cas9 variant. Cytotherapy 2018;20(5):S18.
[31] Bettens K, Sleegers K, Van Broeckhoven C. Genetic insights in alzheimer's disease. Lancet Neurol 2013;12:92–104.
[32] Schellenberg GD, Bird TD, Wijsman EM, Orr HT, Anderson L, et al. Genetic linkage evidence for a familial Alzheimer's disease locus on chromosome 14. Science 1992;258:668–71.
[33] Levy-Lahad E, Wijsman EM, Nemens E, Anderson L, Goddard KA, et al. A familial Alzheimer's disease locus on chromosome 1. Science 1995;269:970–3.
[34] Vetrivel KS, Zhang YW, Xu H, Thinakaran G. Pathological and physiological functions of presenilins. Mol Neurodegener 2006;1:4.
[35] Ortiz-Virumbrales M, Moreno CL, Kruglikov I, Marazuela P, Sproul A, et al. CRISPR/Cas9-correctable mutation-related molecular and physiological phenotypes in iPSC-derived Alzheimer's PSEN2 (N141I) neurons. Acta NeuropatholCommun 2017;5:77.
[36] Pires C, Schmid B, Petræus C, Poon A, Nimsanor N, et al. Generation of a gene-corrected isogenic control cell line from an Alzheimer's disease patient iPSC line carrying a A79V mutation in PSEN1. Stem Cell Res 2016;17:285–8.
[37] Poon A, Schmid B, Pires C, Nielsen TT, Hjermind LE, et al. Generation of a gene-corrected isogenic control hiPSC line derived from a familial Alzheimer's disease patient carrying a L150P mutation in presenilin 1. Stem Cell Res 2016;17:466–9.
[38] György B, Lööv C, Zaborowski M, Takeda S, Kleinstiver B, et al. CRISPR/Cas9 mediated disruption of the swedish APP allele as a therapeutic approach for early-onset alzheimer's disease. Mol Ther Nucleic Acids 2018;11:429–40.
[39] Mullan M, Crawford F, Axelman K, Houlden H, Lilius L, et al. A pathogenic mutation for probable Alzheimer's disease in the APP gene at the N-terminus of beta-amyloid. Nat Genet 1992;1:345–7.
[40] Das U, Scott DA, Ganguly A, Koo EH, Tang Y, et al. Activity-induced convergence of APP and BACE-1 in acidic microdomains via an endocytosis-dependent pathway. Neuron 2013;79:447–60.
[41] Genetics Eisenstein M. Finding risk factors. Nature 2011;475:S20–2.
[42] Weisgraber KH, Rall Jr SC, Mahley RW. Human E apoprotein heterogeneity. Cysteine-arginine interchanges in the amino acid sequence of the apo-E isoforms. J Biol Chem 1981;256:9077–83.
[43] Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and alzheimer disease meta-analysis consortium. JAMA 1997;278:1349–56.
[44] Saunders AM, Strittmatter WJ, Schmechel D, George-Hyslop PH, Pericak-Vance MA, et al. Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. Neurol 1993;43:1467–72.
[45] Kanekiyo T, Xu H, Bu G. ApoE and Aβ in Alzheimer's disease: accidental encounters or partners? Neuron 2014;81:740–54.
[46] Wang C, Najm R, Xu Q, Jeong DE, Walker D, et al. Gain of toxic apolipoprotein E4 effects in human iPSC-derived neurons is ameliorated by a small-molecule structure corrector. Nat Med 2018;24:647–57.
[47] Dong LM, Weisgraber KH. Human apolipoprotein E4 domain interaction. Arginine 61 and glutamic acid 255 interact to direct the preference for very low density lipoproteins. J Biol Chem 1996;271:19053–7.
[48] Moreno CL, Della Guardia L, Shnyder V, et al. iPSCderived familial Alzheimer's PSEN2 N141I cholinergic neurons exhibit mutation-dependent molecular pathology corrected by insulin signaling. Mol Neurodegener 2018;13(1):33.
[49] Park H, Oh J, Shim G, et al. In vivo neuronal gene editing via CRISPR-Cas9 amphiphilic nanocomplexes alleviates deficits in mouse models of Alzheimer'sdisease. Nat Neurosci 2019;22(4):524–8.
[50] Tan D, Yao S, Ittner A, et al. Generation of a new tau knockout (tau Dex1) line using CRISPR/Cas9 genome editing in mice. J Alzheimers Dis 2018;62(2):571–8.
[51] Paquet D, Kwart D, Chen A, et al. Efficient introductionof specific homozygous and heterozygousmutations using CRISPR/Cas9. Nature 2016;533(7601):125–9.
[52] Fang B, Jia L, Jia J. Chinese presenilin-1 V97Lmutation enhanced Ab42 levels in SH-SY5Y neuroblastomacells. Neurosci Lett 2006;406(1–2):33–7.
[53] Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 2016;533(7603):420–4.
[54] Xu TH, Yan Y, Kang Y, Jiang Y, Melcher K, Xu HE. Alzheimer's disease-associated mutations increaseamyloid precursor protein resistance to c-secretasecleavage and the Ab42/Ab40 ratio. Cell Discov 2016;2(1):1–4.
[55] Sud hanshu P. Raikwar, Ramasamy Thangavel, Iuliia Dubova, Govindhasamy Pushpavathi Selvakumar, Mohammad Ejaz Ahmed, Duraisamy Kempuraj, Smita A. Zaheer, Shankar S. Iyer, Asgar Zaheer, Targeted gene editing of glia maturation factor in microglia: a novel Alzheimer's disease therapeutic target, Mol. Neurobiol. 56 (2019) 378–393.
[56] Nagata Kenichi, Takahashi Mika, Matsuba Yukio, Okuyama-Uchimura Fumi, Sato Kaori, Hashimoto Shoko, et al. Generation of App knock-in mice reveals deletion mutations protective against Alzheimer's diseaselike pathology. Nat Commun 2018;9:1800.
[57] Wong Eitan, Liao George P, Chang Jerry C, Peng Xu, Li Yue-Ming, Greengard Paul. GSAP modulates γ-secretase specificity by inducing conformational change in PS1. Proc Natl Acad Sci 2019;116:6385–90.
[58] Zhang Yun-wu, Thompson Robert, Zhang Han, Huaxi Xu. APP processing in Alzheimer's disease. Mol Brain 2011;4:3.
[59] Wadhwani Anil R, Affaneh Amira, Van Gulden Stephanie, Kessler John A. Neuronal apolipoprotein E4 increases cell death and phosphorylated tau release in alzheimer disease. Ann Neurol 2019;85:726–39.
[60] Sun W, Ji W, Hall JM, et al. Self-assembled DNA nanoclews for the efficient delivery of CRISPR-Cas9 for genome editing. Angew Chem Int Ed 2015;54:12029–333.
[61] Aghamiri S, Talaei S, Ghavidel AA, et al. Nanoparticles-mediated CRISPR/Cas9 delivery: recent advances in cancer treatment. J Drug Deliv Sci Technol 2020;56:101533.
[62] Kulkarni JA, Cullis PR, Van Der Meel R. Lipid nanoparticles enabling gene therapies: from concepts to clinical utility. Nucleic Acid Ther 2018;28(3):146–57.
[63] Wang D, Zhang F, Gao G. CRISPR-based therapeutic genome editing: strategies and in vivo delivery byAAV vectors. Cell 2020;181(1):136–50.
[64] Anzalone AV, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. Nature 2019;576:149–57.
[65] Veres A, Gosis BS, Ding Q, Collins R, Ragavendran A, Brand H, et al. Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. Cell Stem Cell 2014;15:27–30.
[66] Soldner F, Stelzer Y, Shivalila CS, Abraham BJ, Latourelle JC, Barrasa MI, et al. Parkinson-associated risk variant in distal enhancer of α-synuclein modulates target gene expression. Nature 2016;533:95.
[67] Arias-Fuenzalida J, Jarazo J, Qing X, Walter J, Gomez-Giro G, Nickels SL, et al. FACS-assisted CRISPR-Cas9 genome editing facilitates Parkinson's disease modeling. Stem Cell Rep 2017;9:1423–31.
[68] Qing X, Walter J, Jarazo J, Arias-Fuenzalida J, Hillje AL, Schwamborn JC. CRISPR/cas9 and piggyBac-mediated footprintfree LRRK2-G2019S knock-in reveals neuronal complexity phenotypes and αSynuclein modulation in dopaminergic neurons. Stem Cell Res 2017;24:44–50.
[69] Potting C, Crochemore C, Moretti F, Nigsch F, Schmidt I, Manneville C, et al. Genome-wide CRISPR screen for PARKIN regulators reveals transcriptional repression as a determinant of mitophagy. Proc Natl Acad Sci U S A 2018;115:E180–9.
[70] Song C, Charli A, Luo J, Riaz Z, Jin H, Anantharam V, et al. Mechanistic interplay between autophagy and apoptotic signaling in Endosulfaninduced dopaminergic neurotoxicity: relevance to the adverse outcome pathway in pesticide neurotoxicity. Toxicol Sci 2019;169:333–52.
[71] Gordon R, Neal ML, Luo J, Langley MR, Harischandra DS, Panicker N, et al. Prokineticin-2 upregulation during neuronal injury mediates a compensatory protective response against dopaminergic neuronal degeneration. Nat Commun 2016;7:12932.
[72] Kempuraj D, Khan MM, Thangavel R, Xiong Z, Yang E, Zaheer A. Glia maturation factor induces interleukin-33 release from astrocytes: implications for neurodegenerative diseases. J NeuroImmune Pharmacol 2013;8:643–50.
[73] Selvakumar GP, Ahmed ME, Raikwar SP, Thangavel R, Kempuraj D, Dubova I, et al. CRISPR/Cas9 editing of glia maturation factor regulates mitochondrial dynamics by attenuation of the NRF2/HO-1 dependent ferritin activation in glial cells. J NeuroImmune Pharmacol 2019:1–14.
[74] Yang W, Tu Z, Sun Q, Li X. CRISPR/Cas9: implications for modeling and therapy of neurodegenerative diseases. Front Mol Neurosci 2016;9:30.
[75] Chen Y, Dolt KS, Kriek M, Baker T, Downey P, Drummond NJ, et al. Engineering synucleinopathy-resistant human dopaminergic neurons by CRISPR-mediated deletion of the SNCA gene. Eur J Neurosci 2019;49:510–24.
[76] Park H, Oh J, Shim G, Cho B, Chang Y, Kim S, et al. In vivo neuronal gene editing via CRISPR–Cas9 amphiphilic nanocomplexes alleviates deficits in mouse models of Alzheimer's disease. Nat Neurosci 2019;22:524.
[77] Monteys AM, Ebanks SA, Keiser MS, Davidson BL. CRISPR/Cas9 editing of the mutant huntingtin allele in vitro and in vivo. Mol Ther 2017;25:12–23.
[78] Tang T, Tu H, Chan EYW, Maximov A, Wang Z, Wellington CL, et al. Huntingtin and huntingtin-associated protein 1 influence neuronal calcium signaling mediated by inositol-(1, 4, 5) triphosphate receptor type 1. Neuron 2003;39:227–39.
[79] Wu J, Ryskamp D, Birnbaumer L, Bezprozvanny I. Inhibition of TRPC1-dependent store-operated calcium entry improves synaptic stability and motor performance in a mouse model of Huntington's disease. J Huntingt Dis 2018;7:35–50.
[80] Lee J, Park EH, Couture G, Harvey I, Garneau P, Pelletier J. An upstream open reading frame impedes translation of the huntingtin gene. Nucleic Acids Res 2002;30:5110–9.
[81] Kolli N, Ming L, Maiti P, Rossignol J, Dunbar GL. CRISPRCas9 mediated gene-silencing of the mutant huntingtin gene in an in vitro model of Huntington's disease. Int J Mol Sci 2017;18:754.
[82] Kantor B, Tagliafierro L, Gu J, Zamora ME, Ilich E, Grenier C, Huang ZY, Murphy S, Chiba-Falek O. Downregulation of SNCA Expression by Targeted Editing of

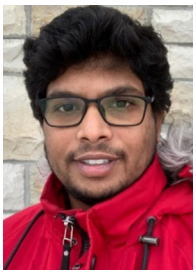
DNA Methylation: A Potential Strategy for Precision Therapy in PD. Mol Ther. 2018 Nov 7;26(11):2638-2649. doi: 10.1016/j.ymthe.2018.08.019. Epub 2018 Aug 29. PMID: 30266652; PMCID: PMC6224806.
[83] Zhou Xiaoqing, Xin Jige, Fan Nana, Zou Qingjian, Huang Jiao, Zhen Ouyang Yu, et al. Generation of CRISPR/Cas9-mediated gene-targeted pigs via somatic cell nuclear transfer. Cell Mol Life Sci 2015;72:1175–84.
[84] Arias-Fuenzalida Jonathan, Jarazo Javier, Qing Xiaobing, Walter Jonas, Gomez-Giro Gemma. Sarah Louise Nickels, Holm Zaehres, Hans Robert Schöler, Jens Christian Schwamborn, FACS-assisted CRISPR-Cas9 genome editing facilitates Parkinson's disease modeling. Stem Cell Rep 2017;9:1423–31.
[85] Wang Xianlong, Cao Chunwei, Huang Jiaojiao, Yao Jing, Hai Tang, Zheng Qiantao, et al. One-step generation of triple gene-targeted pigs using CRISPR/Cas9 system. Sci Rep 2016;6:20620.
[86] Yang Su, Chang Renbao, Yang Huiming, Zhao Ting, Hong Yan. Ha Eun Kong, Xiaobo Sun, Zhaohui Qin, Peng Jin, Shihua Li, CRISPR/Cas9-mediated gene editing ameliorates neurotoxicity in mouse model of Huntington's disease. J Clin Invest 2017;127:2719–24.
[87] Jun Wan Shin, Kyung-Hee Kim, Michael J. Chao, Ranjit S. Atwal, Tammy Gillis, Marcy E. MacDonald, James F. Gusella, Jong-Min Lee, Permanent inactivation of Huntington's disease mutation by personalized allele-specific CRISPR/Cas9, Hum.Mol. Genet. 25(2016) 4566-4576.
[88] Jun Wu, Ryskamp Daniel, Birnbaumer Lutz, Bezprozvanny Ilya. Inhibition of TRPC1-dependent store-operated calcium entry improves synaptic stability and motor performance in a mouse model of Huntington's disease. J Huntingt Dis 2018;7:35–50.
[89] Nivya Kolli Lu, Ming Panchanan Maiti, Rossignol Julien, Dunbar Gary L. CRISPRCas9 mediated gene-silencing of the mutant huntingtin gene in an in vitro model of Huntington's disease. Int J Mol Sci 2017;18:754.
[90] Xu CL, Park KS, Tsang SH. CRISPR/Cas9 genome surgery for retinal diseases. Drug Discov Today Technol 2018;28:23–32.
[91] DiCarlo JE, Sengillo JD, Justus S, Cabral T, Tsang SH, Mahajan VB. CRISPR-Cas Genome Surgery in Ophthalmology. Transl Vis Sci Technol 2017;6(3):13.
[92] Karimian A, Gorjizadeh N, Alemi F, Asemi Z, Azizian K, Soleimanpour J, et al. CRISPR/Cas9 novel therapeutic road for the treatment of neurodegenerative diseases. Life Sci 2020;259:118165.
[93] Xu T, Li L, Liu YC, Cao W, Chen JS, Hu S, et al. CRISPR/Cas9-related technologies in liver diseases: from feasibility to future diversity. Int J Biol Sci 2020;16(13):2283–95.
[94] Schneller JL, Lee CM, Bao G, Venditti CP. Genome editing for inborn errors of metabolism: advancing towards the clinic. BMC Med 2017;15(1):43.
[95] German DM, Mitalipov S, Mishra A, Kaul S. Therapeutic Genome Editing in Cardiovascular Diseases. JACC Basic Transl Sci 2019;4(1):122–31.
[96] Mirza Z, Karim S. Advancements in CRISPR/Cas9 technology-Focusing on cancer therapeutics and beyond. Semin Cell Dev Biol 2019;96:13–21.
[97] Kuai R, Yuan W, Li W, Qin Y, Tang J, Yuan M, et al. Targeted delivery of cargoes into a murine solid tumor by a cell-penetrating peptide and cleavable poly (ethylene glycol) comodified liposomal delivery system via systemic administration. Mol Pharm 2011;8:2151–61.
[98] LeWitt PA, Rezai AR, Leehey MA, Ojemann SG, Flaherty AW, Eskandar EN, et al. AAV2-GAD gene therapy for advanced Parkinson's disease: A double-blind, sham-surgery controlled, randomised trial. Lancet Neurol 2011;10:309–19.



Shanu Bhardwaj is currently pursuing Master of Science (M.Sc.) in Biotechnology from Delhi Technological University (DTU), India. Her area of interest includes Neurodegeneration, Neuroinformatics, and Neurotherapeutics. She has already published articles in reputed journals including Open Biology and Journal of Molecular Neuroscience.



Dr. Kavindra Kumar Kesari is currently working as Senior Researcher in the Department of Applied Physics at the Aalto University, Espoo, Finland. He obtained Doctoral, Master, and Bachelor degreés in Biotechnology from India. He served as an Assistant Professor position at Jaipur National University, Jaipur, India. He obtained his first Postdoctoral research from University of Eastern Finland, Kuopio, Finland. He has published over 70 papers in reputed journals, 20 book chapters, 5 books and presented over 25 papers at national and international scientific meetings. His Hirsch index (h-index) as of June 2021 is 29 on Google Scholar, 24 on Scopus. He has received four times young scientist award. His research interest focuses on neurobiology, radiation and cancer biology and therapeutic research.



Mahesh Rachamalla is currently pursuing PhD studies at Department of Biology, University of Saskatchewan, Canada. He received his master's degree in Pharmacology and Toxicology in 2011 from National Institute of Pharmaceutical Education and Research (NIPER), Mohali. He has published 5 research articles, 2 review articles and 1 book chapter. His area of research is majorly focused on investigating molecular mechanisms involved in neurotoxicity and cognitive impairment upon exposure to metals.



Dr. Shalini Mani is Associate Professor in the Department of Biotechnology at Jaypee Institute of Information Technology, Noida, India. She has total 17 years of research and 10 years of teaching experience. She completed her PhD from Centre for Cellular and Molecular Biology, Hyderabad, where she worked on detailed analysis of clinical biochemical and genetic analysis of Leigh syndrome patients. She also received a prestigious Endeavour Research fellowship award from Australia and did her post-doctorate from Murdoch Children Research Institute, Melbourne Australia. A resurgence of interest in the study of mitochondria during the past decades has revealed that mitochondria also play key roles in cell signaling, proliferation, cell metabolism and cell death, and that genetic and/or metabolic alterations in mitochondria contribute to a number of diseases, including cancer, diabetes and neurological disorders. Thus, as a part of fundamental research, her research interest primarily focuses on studying the role of mitochondria in different chronic diseases as well as effect of different environmental factors on mitochondrial metabolism. As a translational researcher, she is also exploring the molecular mechanism for therapeutic benefits of some selected herbs on mitochondrial activity in different diseases conditions such as diabetes, cancer and neurological disorders.



Dr. Ghulam Md Ashraf is working as Associate Professor in King Fahd Medical Research Center, King Abdulaziz University, Saudi Arabia. He has a PhD degree in Biochemistry with over ten years of teaching and research experience in various disciplines of biological and medical sciences. His primary fields of research are biochemistry and neurology, currently focusing on understanding the molecular and behavioral mechanism of effects of anti-diabetic drugs in pyschotic and dementia conditions. Dr. Ashraf's lab is currently investigating the effect of anti-diabetic drugs in dementia, and their role in possible attenuation/reversal of anti-psychotic drugs induced weight gain in psychotic conditions. He is also investigating molecular and behavioral mechanisms of novel therapeutic combinations in multiple sclerosis and Alzheimer's disease. Dr. Ashraf has published 284 research articles (citations: 5213, H-index: 35, i10 index: 135) and have been involved in 21 research grants [14 ongoing (10 PI, 4 CoI) and 7 completed (4 PI, 3 CoI)]. He is also involved in editorial roles in renowned scientific journals like Scientific Reports (Senior EBM), PLoS One (Academic Editor), Current Neuropharmacology (Section Editor). Dr. Ashraf is professionally associated with the Royal Society of Medicine (Fellow), Royal Society of Biology (Member) and American Society for Biochemistry and Molecular Biology (Member).



Dr. Saurabh Kumar Jha is working as an Assistant Professor in the School of Engineering and Technology (Department of Biotechnology) at Sharda University, Greater Noida, India. He obtained his M.Sc. degree in Biotechnology from VIT University, Vellore, and Ph.D. degree in Biotechnology was awarded from Delhi Technological University, India. His areas of interest and expertise include Stem cell, cell signaling and disease. He has published 34 papers in peer-reviewed international journals such as Journal of Neurochemistry, Journal of Alzheimer's disease and BBA-Molecular Basis of Disease.



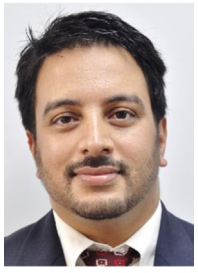
Dr. Pravir Kumar is working as a Professor and Head of the Department Biotechnology at Delhi Technological University, Delhi, India. Before joining DTU, Prof. Kumar served as an Associate Professor at Vellore Institute of Technology University, Vellore, India. He has obtained MS degree from BHU, Varanasi, India with Molecular and clinical genetics specialization, and PhD degree from J. W. Goethe University, Germany in the field of coronary artery diseases and cardiovascular physiology. Before returning to India, he served in the Neurology Department at Tufts University School of Medicine, Boston, USA as a postdoctoral fellow and later at faculty position. He has published more than 75 papers in the peer reviewed high impact factor international journals (h-index: 22; i-10: 29; Citation: 2319; cumulative IF: 200). His areas of research interest and expertise include protein aggregation, molecular chaperone and ubiquitin E3 ligase in neurodegenerative disorders along with the aberrant cell cycle re-entry into aged neurons and muscles. He is an editorial board member in the prestigious Scientific Reports (Nature publications) and reviewer of reputed publication houses like Elsevier, Springer, BMC, Bentham, Oxford among others. He was also one of the Deans of DTU.



Dr. Rashmi K Ambasta has finished her Master's degree from the Department of Zoology, Banaras Hindu University, India. She obtained her PhD degree from J. W. Goethe University in Frankfurt, Germany. She has finished her PhD in 2004 and joined Boston, MA, USA and did two Postdoctoral training sponsored by NIH. She joined the faculty position at VIT, Vellore, India as an Associate Professor. She has published more than 60 papers in the peer reviewed high impact international journals (h-index: 19; i-10: 25; Citation: 1710; cumulative IF: 160). Dr. Ambasta was awarded with SERB Young Scientist in 2012, Delhi followed by CSIR Scientific Pool Officer award in 2016 at DTU, Delhi. During her 20 years of research and teaching experience she has been awarded with some prestigious awards like CSIR-JRF, ICMR-JRF, ASRB-LS, Centenary IISC-Bangalore Postdoctoral Fellow, Elsevier Outstanding Reviewer Award and Commendable Research Award 2017, 2019, 2020 in DTU. Her research interest is in the field of Cancer Drug Targeting and Neuro-oncology, Stem cell and Diabetes. Currently, she is working as a guest faculty in the department of Biotechnology at Delhi Technological University, Delhi, India.



Dr. Harish Dureja is presently working as a Professor in Pharmaceutics, Department of Pharmaceutical Sciences, M. D. University, Rohtak. He is also holding the position of Director, Centre for IPR Studies and Director, Professional Consultancy Services at M. D. University, Rohtak. He has a total of 21 years teaching and research experience. He has filed one Indian patent, published twenty book chapters and 210 publications in various international and national Journals of repute. He has been awarded research projects of worth 1.2 Crore. He has also delivered more than 150 invited lectures in various conferences/seminars and symposia. He has guided 60 PG scholars and 10 Ph.D. students.



Dr. Hari Prasad Devkota completed his Bachelor of Pharmaceutical Sciences (B. Pharm.) from Pokhara University, Nepal in 2005. After working at the same institute as Teaching Associate till 2007, he received Japanese Government (MEXT) Scholarship and entered Graduate School of Pharmaceutical Sciences, Kumamoto University, Japan in 2007 and completed Master and PhD in Pharmaceutical Sciences in 2010 and 2013, respectively. He was working as Postdoctoral Fellow in Kumamoto University (2013-2014) with the support from Takeda Science Foundation (TSF), Japan. From 2014, he is working as Assistant Professor at HIGO Program, Kumamoto University. His main research interests are plant-based bioactive natural products, functional foods and ethnopharmacology. Till now, he has authored more than 130 publications including original research articles, review articles, books and book chapters. He is actively involved in the research on medicinal plants from Asia and Africa for their ethnopharmacological surveys, chemical constituents and pharmacological activities.



Dr. Gaurav Gupta holds a doctoral degree from Pacific University, Udaipur, Rajasthan, India. He has more than nine years of experience in molecular and biochemical pharmacology including phytochemistry, respiratory diseases, psychopharmacology, and cancer biology by employing experimental animal models to understand the cellular and molecular mechanism. Dr. Gupta is dedicated to improving outcomes in healthcare through initiatives in pharmacology and phytochemistry research and effective teaching in the field of Pharmaceutical Sciences. Dr. Gupta is currently associated as an Associate Professor with Suresh Gyan Vihar University, Jaipur, Rajasthan, India. He has >200 scientific articles in the national and international journals of repute. The total cumulative impact factor of his published papers is around 300 (as per Thomson Reuters).



Dr. Dinesh Kumar Chellappan holds PhD and Master's degree in Pharmaceutical Sciences from Manipal University, India. He worked on several research projects focused on inflammatory diseases with the core area in diabetes treatment and management. He is currently actively involved in undertaking research in the area of chronic inflammatory pulmonary diseases, particularly, developing and testing novel drug delivery systems. He has published >100 publications in reputed national and international journals with cumulative impact factor of >200. He has presented several papers at conferences and have won best poster and best oral awards.



Dr. Sachin Kumar Singh is a Pharmaceutical Analyst and holds Ph.D degree in Pharmaceutical Sciences from Karpagam University, Coimbatore, India. Currently he is working as Professor, School of Pharmaceutical Sciences, Lovely Professional University (LPU), India. He has completed four industrial projects. Dr. Singh's research work on polypeptide-k has got recognition from Punjab State Council for Science and Technology and DST/SERB. He has received a grant of 2.43 million under FastTrack Young Scientists Scheme for the development of SNEDDS of polypeptide-K for the effective treatment of diabetes mellitus. He has published more than 120 research papers in various national and international journals with cumulative impact factor 250 and h-index 21, filed ten patents and 10 international book chapters. He has supervised 30 M. Pharm projects and 8 PhD thesis and currently guiding 3 M. Pharm projects and 6 PhD thesis. He is reviewer of more than 40 international journals indexed in Web of Science and Scopus. Dr. Singh is a recipient of young scientist award, 2013 (DST/SERB govt of India), Bharat Shiksha Ratan Award, 2014 (G.S.H.E.G., New Delhi, India) and Dr. B. C. Deb Memorial Award, 2020 for popularization of science (Indian Science Congress Association, DST, India).



Dr. Kamal Dua, Senior lecturer, Discipline of Pharmacy, Graduate School of Health, University of Technology Sydney (UTS), has a research experience of over 12 years working in the field of drug delivery targeting inflammatory diseases. Dr. Dua is also a Node Leader of Drug Delivery Research in the Centre for Inflammation at Centenary Institute/UTS, where the targets identified from the research projects are pursued to develop novel formulations as the first step towards translation into clinics. Dr. Dua researches in two complementary areas; drug delivery and immunology, specifically addressing how these disciplines can advance one another helping the community to live longer and healthier. This is evidenced by his extensive publication record in reputed journals. Dr. Dua's research interests focus on harnessing the pharmaceutical potential of modulating critical regulators such as Interleukins and microRNAs and developing new and effective drug delivery formulations for the management of inflammation in chronic airway diseases and cancer.



Dr. Janne Ruokolainen is a Professor in the Department of Applied Physics, Aalto University, Espoo, Finland. He gained his PhD (Tech) in Materials Physics from Helsinki University of Technology and Docentship in polymer physics and soft materials microscopy. His current research activities focus the characterization of molecular bioactive compounds extracted from plant species in neurodegenerative disease treatment. He has published over 250 scientific papers in peer reviewed and refereed journals. Since 1995, he has gained high bibliometric factor, where H-index 62 (google scholar). His awards and research honours include invitation to the Alfred Nobel Symposium (2005), Young Scientist Award by Foundation of Technology (2002), Prize for the best PhD Thesis in the field of technical sciences by Tekniikan Akateemisten Liitto and Tekniska Föreningen in Finland (TEK/TFiF) (2001), IUPAC Prize for Young Chemists – Honourable Mention Award (2001).



Dr. Mohammad Amjad Kamal is a Distinguish Adjunct Professor at King Fahd Medical Research Center (KFMRC), King Abdulaziz University, Jeddah, Saudi Arabia, where he was a full Professor from March 2011 to March 2013. Prof. Kamal's overall biochemical research output had culminated in more than 550 publications in internationally respected journals. His research was pivotal in supporting the development of the novel anti-Alzheimer's agents from the laboratory to the clinic via collaboration with Dr. Nigel H. Greig (Chief of Laboratory of Neurosciences and Gerontology department in the USA). Moreover, he is serving as a regional/editorial member/guest editor of several reputed scientific journals - this is undertaken on an honourable basis.



Dr. Shreesh Ojha is currently Associate Professor in the Department of Pharmacology and Therapeutics at the College of Medicine and Health Sciences, United Arab Emirates University, Abu Dhabi, United Arab Emirates. He holds a doctoral degree in Pharmacology from All India Institute of Medical Sciences, India following postdoctoral training from College of Medicine and BIO 5 Institutes of University of Arizona, Tucson, Arizona, USA. He also worked as Research Scientist in Clinical Pharmacology and Drug Development division with a Pharmaceutical company, Daichii Sankyo, Japan formerly known as Ranbaxy Research Laboratories. His expertise span from preclinical to clinical pharmacology across industries to academia. He is on the editorial board of many journals and reviewer for several reputed journals. Dr. Ojha has published more than 150 manuscripts including original research articles and reviews. His H-index is 39 with more than 4700 citations (google scholar). His research interest is natural products pharmacology and caanabinoid signaling in neurodegenerative and cardiovascular diseases. He is reciepient of many research grants including Sheikh Hamdan Award for Medical Sciences and Abu Dhabi Department of Education and Knowledge (ADEK), United Arab Emirates.



Dr. Niraj Kumar Jha is currently Assistant Professor in the Department of Biotechnology, Sharda University, Greater Noida, India. He is a former Assistant Professor of the Faculty of Biotechnology, Noida Institute of Engineering and Technology (NIET), affiliated with Abdul Kalam Technical University (AKTU), India. He holds a Doctorate in Biotechnology from Delhi Technological University and a postgraduate degree in Biotechnology from Vellore Institute of Technology, TN, India. During the doctorate program, DBT granted him a fellowship as a financial boost to accomplish research work. He has published over 60 quality scholarly work and 5 book chapters in prestigious international journals and presented over 10 research papers in top national and international conferences. He is an editorial board member of various high repute international journals and assisted significantly in reviewing articles of many international journals such as Scientific reports, Life sciences, European journal of pharmaceutical sciences and Journal of food biochemistry. Currently, he is also serving as a guest editor for various top journals of international repute including Antioxidants, Environmental science and pollution research and Immuno. He has executed various research projects and supervised several doctoral scholars, masters, and bachelor students. His vast research expertise covers the areas of Natural products, Cell signaling and Disease.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/353707234

Customer Relationship Management of Different Companies

Conference Paper · August 2021

DOI: 10.5281/zenodo.5163054

CITATIONS
0

READS
20

2 authors:



Mozammil Hassan
Delhi Technological University
4 PUBLICATIONS **0** CITATIONS
SEE PROFILE



Sadique Hassan
Delhi Technological University
1 PUBLICATION **0** CITATIONS
SEE PROFILE

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

Project Recent Advancements in Supply Chain Management View project

Project Customer Relationship Management of Different Companies View project

All content following this page was uploaded by Mozammil Hassan on 05 August 2021.

The user has requested enhancement of the downloaded file.

HBRP PUBLICATION

Recent Trends in Production Engineering
Volume 4 Issue 2
DOI: https://doi.org/10.5281/zenodo.5163054

Customer Relationship Management of Different Companies

Mozammil Hassan[1]*, Sadique Hassan[2]
[1]Department of Mechanical Engineering, Delhi Technological University
[2]Department of Environmental Engineering, Delhi Technological University

[*]Corresponding Author
E-Mail Id:-mozammilhassan_2k18me129@dtu.ac.in

ABSTRACT

In today's fiercely competitive market, customers are at the centre of all business transactions in every sector. Customers are the main reason for every business organization's existence in today's global market. Furthermore, achieving global competitiveness and long-term success requires a strategic focus on customer satisfaction. As a result, in this rising economy, the concept of Customer Relationship Management has become a prominent corporate approach. These improvements have altered producers' consciousness from conventional advertising to fashionable advertising, which contains a good deal greater than simply developing a product, pricing it, selling it, and making it to be had to the goal market. It requires the improvement of trust, a binding force, and a value-brought courting with clients. Customer courting management, or CRM, is the system of organising a cooperative and collaborative courting among the purchaser and seller. Thus, through this article, I have focused on understanding what CRM is and to investigate CRM policies and their impact in several MNCs such as Amazon, BMW, and Apple Inc.

Keywords:-*Customer relationship management, customer satisfaction, competitive advantage, customer information, customer loyalty*

INTRODUCTION

Customers are now at the centre of every business transaction in every sector in today's fiercely competitive market. Customers are the main reason for every business organization's existence in today's global market. The ability to satisfy and retain consumers has become a big concern for service companies all around the world. Managing the relationship with the customer's subsequent actions is an important aspect in achieving competitive advantage. As a result, premier marketers in the twenty-first century are increasingly adopting relationship management methods. One-to-one marketing is now the strategy slogan for achieving global competitiveness and long-term growth sustainability. According to numerous studies, it is more beneficial for a corporation to keep an existing customer than to acquire a new one. Similarly, according to Peter Drucker, a brilliant management guru, "the objective of company is to generate and keep a customer."

It is a set of people, procedures, software, and internet capabilities that enable a company to manage its customer interactions in an efficient and methodical manner. CRM's purpose is to understand and anticipate the demands of present and new customers, as well as to boost customer retention and loyalty while broadening the range of products and services supplied. CRM's goal is to collect data from every encounter a company has with a client, regardless of channel, and save it in a CRM system so that the organisation can properly understand

Recent Trends in Production Engineering
Volume 4 Issue 2
DOI: https://doi.org/10.5281/zenodo.5163054

customer behaviour. CRM software enables a business to create a customer database that all stakeholders, including management, salespeople, customer support providers, and even customers, can use in order to assess the customer's demands in terms of product and offering.

LITERATURE REVIEW

In today's world, the technological revolution has completely transformed and changed the way businesses operate. As a result, technology plays a critical role in establishing more successful client interactions. Customer relationship management (CRM) is a cutting-edge strategy that allows a company to effectively track both existing and potential consumers. It makes a lot more sense to focus on keeping existing customers happy than putting a lot of effort into recruiting new ones. CRM is a machine that aids agencies in growing client relationships and streamlining tactics which will increase sales, enhance client service, and increase profits (Nagi,2005).

The CRM concept is a method of managing customer relationships that evolved from the early concept of relationship marketing (Gummeson, 2008). CRM is also defined as " The values and techniques of Relationship Marketing with a unique emphasis on the connection among a consumer and a supplier where businesses can store customer and prospect data, track customer communications, and share this with similar groups," according to Gummesson (2008).

It aids in the management of client connections as well as the expansion of the company. CRM, on the other hand, can be thought of as a company-wide business strategy aimed at increasing revenue and profitability, lowering costs, and increasing customer loyalty. CRM is a synergistic combination of customer connection marketing methods and information and communication technologies to establish worthwhile and long-time period relationships with clients and different stakeholders in a business. CRM generation is likewise a cross-purposeful integration of people, processes, operations, and advertising strategy supported by ICT applications that enables a business to generate efficient customer value. As a result, CRM's major goal is to upgrade the customer's connection from "stranger" to "acquaintance," then to "friend," and lastly to "partner" (Raizda, 2011). Leena, et al. (2015), on the other hand, saw CRM as a "umbrella idea" that puts the customer at the centre of every firm.

CRM's Advantages

1. Decreased sales costs: Sales costs are reduced as a result of existing customers becoming more responsive. Furthermore, with a better understanding of channels and distributors, partnerships become more effective, and marketing campaign expenses are decreased.
2. Higher client profitability: As the client pockets percentage grows, up-selling, cross-selling, and follow-up income grow, and greater pointers arrive with better client happiness, client profitability will rise.
3. To keep a regular quantity of business, there may be no want to collect as many clients: because the` variety of long-time period clients grows, the want to recruit as many clients reduces.
4. Increased client retention and loyalty: As clients live longer, purchase extra, and purchase extra frequently, client retention rises. In addition, clients are much more likely to take initiative, which strengthens the tie among them and complements client loyalty.

Recent Trends in Production Engineering
Volume 4 Issue 2
DOI: https://doi.org/10.5281/zenodo.5163054

CASE STUDY 1- APPLE INC

OVERVIEW

Apple, the most well-known company in the world of creative technology today, offers a wide range of devices and gadgets, including computers, tablets, smartphones, watches, laptops, and more. Since its inception in 2001, it has centred on retail and its consumers, and has grown to emerge as the third biggest PC producer within side the United States.

As a result, consumer relationships have turn out to be a critical awareness of Apple's increase strategy, and the agency has turn out to be a logo in its personal right, just like its products. Apple has captured its personal marketplace percentage in a quite brief duration of time, and consumer loyalty makes it hard for different manufacturers to go into and compete.

NEED FOR CRM IN APPLE

1. For over a decade, Apple has hired CRM suggestions in its customer support divisions, which includes maintaining in contact with clients who've used their in-save gadgets via emails if you want to gauge how they experience at the same time as speak with Apple employees.
2. Apple has been employing CRM in their retail base to establish a long-term relationship between the company and its customers. Rather than focusing solely on selling their high-end products, their strategy focuses on informing and educating customers about the benefits of having them. Customers are attracted to Apple products because of their features, which entice them to purchase one. Apple has improved the client relationship and turned it into an experience for them.
3. Data accumulating is on the coronary heart of Apple's CRM policy. Customers have to sign in the usage of their precise Apple ID whether or not they buy an Apple tool together with an iPhone, iPad, or use Apple software program together with iTunes. The complete set of information is utilised to help outline marketing and marketing in order that it may be greater exactly centered at a capability customer.
4. Apple has advanced a cloud-primarily based totally CRM answer this is used now no longer simply with the aid of using the business enterprise however additionally with the aid of using its customers. Businesses advantage from Apple's CRM statistics collecting, which Apple also can exploit, growing a pyramid effect.

CRM OF APPLE

1. CRM's purpose is to enhance present and capacity clients' customer support experiences. Apple's Apple Genius tech assist method permits clients to connect to technicians in man or woman for a better dialogue of troubles and solutions.
2. Apple has in reality made extraordinary use of CRM in coping with face-to-face communique and growing emblem awareness. They area a top rate on CRM-primarily based totally marketing and marketing and customized mailings, which has allowed them to create an entire life revel in as opposed to only a transaction.
3. As a success international corporation, the agency want a CRM device that lets in it to speak with clients on a one-to-one basis. This will even function a manual and motivator for small firms trying to enforce comparable forms of interpersonal CRM.

CASE STUDY 2- AMAZON

OVERVIEW

It isn't always an exaggeration to assert that Amazon has had a world-elegance

Recent Trends in Production Engineering
Volume 4 Issue 2
DOI: https://doi.org/10.5281/zenodo.5163054

CRM coverage for over decades. They are the maximum relied on and desired on line purchasing portal, with hundreds of thousands of dependable customers. This is made viable with the aid of using a CRM coverage this is well-managed, efficient, and efficiently framed.

Amazon is famous for supplying a extensive variety of amazing merchandise to customers in a single handy location. The first-rate component is that the agency has controlled to preserve its patron relationships through the years while not having to cope with them face to face.

The suggestions which might be supplied to the patron primarily based totally on his previous buy help the consumer in choosing the proper product and make purchasing simple, convenient, and enjoyable. Amazon presents an easy-to-use interface for its clients with the aid of using saving all in their information, so destiny purchases cross easily and promptly. The innovative use of information makes the consumer experience favoured and special, and it is a unique method to draw clients.

THE REASON FOR CRM'S SUCCESS

Amazon has its very own in-residence software program this is tailor-made to its needs. Their CRM is the using pressure in the back of all in their useful features. Their programme collects records approximately every consumer, including their place and former purchases, and tailors the user's on-web website online enjoy appropriately.

Even earlier than the level of human contact, the CRM assists the consumer in getting solutions to their questions. Customers have complete get entry to to all facts approximately the product they ordered, such as transport instances and dates, transport status, price methods, and go back policies. This notably reduces the call for and fee of customer support personnel.

AMAZON'S CRM APPLICATIONS

1. COLLECTION OF DATA - Customers who store on Amazon need to sign in an account if you want to make destiny transactions extra convenient. Customers may be despatched with gives and promotions primarily based totally on their preceding purchases the use of those accounts, giving Amazon the capacity to goal marketing.
2. PERSONAL DATA STORAGE - When a client creates an account, all in their private facts, account facts, address, and make contact with facts are stored for destiny transactions. Customers experience the convenience of shopping.
3. RECOMMENDATIONS - Amazon was the first to offer product recommendations. When a consumer logs into his account, Amazon suggests a number of things that he might be interested in. This advice is based on an analysis of the customer's previous purchasing habits. It has also added a "customer who bought this" option and other things that appear to be highly promising in terms of increasing sales without putting pressure on customers.
4. CUSTOMER SUPPORT - Amazon's go back coverage is completely online. When a customer has a hassle with a product, she or he contacts a customer support representative, who has whole get entry to the customer's buy records and different information. This enables the go back coverage and guarantees that any troubles are resolved right away and fast.
5. KINDLE MARKETPLACE - The kindle merchandise and enjoy aren't the equal without CRM due to all of the specifics approximately non-public accounts, storage, payments, and recommendations.

Recent Trends in Production Engineering
Volume 4 Issue 2
DOI: https://doi.org/10.5281/zenodo.5163054

CASE STUDY 3- BMW

OVERVIEW

BMW is one of the maximum famous manufacturer's withinside the car industry, and its quality, design, and exclusivity are exceptionally valuable via way of means of customers. They are a maker of luxury cars and other vehicles for a very specific market group, and their products have a high level of client loyalty. Relationship management is crucial for BMW since it caters to a specific niche of brand-conscious clients. As a result, maintaining customer interactions and collecting customer information is critical for the organization.

BMW and its CRM

Because BMW caters to a particular demographic, it's far essential to have CRM software program with a view to useful resource with inside the promoting of its merchandise and accomplishing its goal audience, specially while new merchandise are released.

They have been diligently working on their direct marketing strategies and making earnest efforts to improve the efficiency of their loyalty programmers. They are attempting to improve their online presence by utilizing social media portals. As a result, BMW has been trying to set up a CRM coverage in an effort to raise and growth its tries to emerge as a extra famous and a hit emblem with inside the future.

SELECTING A SUCCESSFUL CRM STRATEGY

BMW has recently made efforts to improve its CRM policy. They've turned their concept into a package that collects and stores client data while also determining what exactly their customers want from the business.

CRM is often used by industry organizations and companies to provide information about new products and services, in addition to introducing new technologies so that customers can view promotional products. In addition, BMW also communicates with customers through social media.

As a result, we may conclude that BMW has made excellent use of consumer data to learn more about what customers want from the company.

CONCLUSION

CRM is a technology that aids in the tracking of changing customer profiles, the incorporation of this information into product design, and the addition of value to the particular customer. Customers nowadays are international and feature a excessive call for cognition, recognition, approvals, and respect. They favour marketers who can provide a high-quality product, provide prompt service, repair, resolve issues, and improve products. Marketers want to establish a deep and long-lasting relationship with their customers.

REFERENCES

1. Baran, R. J., & Galka, R. J. (2013). *CRM: the foundation of contemporary marketing strategy*. Routledge.
2. Dyche, J., & O'Brien, M. (2002). *The CRM handbook: A business guide to customer relationship management*. Addison-Wesley Professional.
3. Kostojohn, S., Paulen, B., & Johnson, M. (2011). *CRM fundamentals*. Apress.
4. http://crmsystems.expertmarket.co.uk/BMW-CRM-Case-Study
5. https://www.channelreply.com/blog/view/amazon-crm-strategy

Recent Trends in Production Engineering
Volume 4 Issue 2
DOI: https://doi.org/10.5281/zenodo.5163054

6. http://www.managementparadise.com/forums/marketingmanagement/213123-customer-relationship-management-apple-inc.html.

Cite as

Mozammil Hassan, & Sadique Hassan. (2021). Customer Relationship Management of Different Companies. Recent Trends in Production Engineering, 4(2), 1–6. https://doi.org/10.5281/zenodo.5163054

View publication stats

Axiomathes
https://doi.org/10.1007/s10516-021-09585-0

Defining the Undefined: ⊕

Aman Chaudhary[1] · **Luckshay Batra**[2]

Received: 18 March 2021 / Accepted: 1 August 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract
Linked with existence of the almighty, the operation of division by zero which is considered as undefined or indeterminate or infinite sometimes, has been a topic of serious altercation among mathematicians and philosophers for so long. History is evident of the various attempts made to clearly define the algebra of zero, including the idea of division by zero. This includes the evolution of the idea of zero division and various insights from mathematicians like Euler, Craig and more. The realm of the paper contains all such significant attempts and it deals with the idea of the history that wishes to explore the evolution of infinity and undefined. The prime aim of this paper is to clearly distinguish the three concepts of *Undefined, Indeterminate* and *Infinity*, along with the concept of division by zero. It strives to provide a notation to undefined expressions, along with citing justifications and results with the help of unpretentious examples from key branches.

Keywords Undefined · Indeterminate · Infinity · Division by zero · Impending operations on zero

1 Introduction

The fascination to explore infinity has always attracted the human, be its relation with the almighty or the development of new theories in physics, there is hardly any section not dealing with the notion of infiniteness. Mathematicians and philosophers of various golden centuries have spent quite a fraction of their lives dealing and plunging into infinity, which was once exclaimed and believed to be the undefined concept. But with it, its antagonist 0 has also played a significant role in shaping the mind of the society towards infinity and many theories, co-relating everything that existed. With zero, came

✉ Luckshay Batra
luckshay25@gmail.com

Aman Chaudhary
chaudharyamanhrc@gmail.com

[1] Hansraj College, University of Delhi, Mahatma Hans Raj Marg, Malka Ganj, Delhi, India

[2] Delhi Technological University, Rohini, Delhi, India

Published online: 23 August 2021

the algebra and the idea of division by zero, which is now regarded as undefined and the limit form $\frac{0}{0}$ as indeterminate. From once being '*absolute nothing*' (Romig 1924) to being undefined, there has been a major shift in the thought, with $\frac{0}{0}$ always associated in discussions. To relive the concept of undefined and division by zero, the paper walks through various centuries dealing with the contributions of *Brahmagupta, Bhaskara, Euler* and many more.

No matter how much time we all have spent learning mathematics, it always seem a fine line amongst undefined, indeterminate and infinity, and the idea to resolve it clearly, is what the paper centralizes around. The preliminary use of the three terms, viz. undefined, indeterminate and infinity, can be broadly classified in the branches of Number Theory and Functional Analysis. Here by Number Theory, we mean the arithmetic operations and algebra of numbers and fractions. While by Analysis, the idea of functions, their limit forms and operations are concerned. The paper details in these two aspects mainly, as the rest of their uses in other branches are inspired from these two ideas of arithmetic and functions only. The paper deals in giving a broad idea to the readers regarding the use of the three terms and their inter-relation (if any possible), like how the class of indeterminate is dissolved in undefined while dealing arithmetically. The central aim of the paper is to classify the above mentioned three concepts with clear definitions, and to express the dire need of defining a symbol for the undefined. Some results are also established from the sight of sequences, for better distinguishing.

This paper is partitioned into three sections. Section 2 deals with the evolution of algebra of zero, along with the attempts and contributions made by mathematicians around the globe in the domain of division by zero and related arithmetic, as marked in the history. Along with introducing the notion of symbolism and vivid encounters of undefined expressions in calculus, euclidean geometry, differential equations, etc., this section also proposes the symbol $\circledcirc$ for the undefined. Section 3 marks the distinction of the three concepts from the definition point of view, along with the discussion on the undefined character of $\frac{1}{0}$. It also features some points of the relation of $\frac{1}{0}$ with Indeterminate. Section 4 addresses the three concepts from the perspective of Number Theory and Functional Analysis, along with some convenient results and examples from sequences. Section 5 concludes the paper.

2 Notion of Undefined and Symbolism

2.1 Initial Attempts in the Past

After the discovery of numbers, the unceasing desire of the humans to puzzle their head, was given a boost. Various mathematicians spent their time working on new results and discoveries in the field of numbers, not just to earn fame, but also in a hope to touch the untouched almighty. A contribution to vivid understanding and ideas were evident from both the Eastern and Western minds of the globe. Some of those are briefed here.

2.1.1 The Contribution by Indian Mathematicians

Being the land of many great mathematical and philosophical minds, it was not possible for a country like India, having given the symbol of zero, to lag behind in plunging into this idea of undefined and the division by zero theory. Given below are some of the major names and their various contributions as mentioned in the history, who have thrown a light on this facet.

Major footsteps by Brahmagupta: Various operations were defined in the algebra of real numbers but the division by zero[1] was too cumbersome. The conceptualisation of the zero arithmetic is believed to be laid down by the Indian mathematician and astronomer of the 7th Century, Brahmagupta, who first introduced 0 in the family of real numbers.

'Brahmagupta's Brahmasphuta-siddhanta' (Brahmagupta 1966) is the first text to discuss the arithmetic of zero, giving the formal definition of zero as: "*positive (dhana) and negative (rna), if they are equal, is zero (kham)*". He gave the algebra about the operations of addition, subtraction and multiplication involving zero. Brahmagupta was of the opinion that every operation must result in a number. He treated 0 as a number, and considered that any finite number divided by zero is also a number, and he called it '*taccheda*', remarking—

> "Positive or negative, divided by zero, is a fraction with zero for denominator. Zero divided by negative or positive is either zero or is expressed by a fraction with zero as numerator and the finite quantity as denominator. *Zero divided by zero is zero*."

The idea of Bhaskara: After Brahmagupta, many mathematicians gave their ideas about the same, like Mahavira suggested that the finite number acting as numerator of *taccheda* is not partitioned, as the dividing number is absolute nothing and therefore *taccheda* has no determinate quotient. Sripati also gave the term '*khahara*' for taccheda, signifying expressions having zero (*kha*) as the divisor.

However the most significant contribution could be made by Bhaskara only, in the 12th century, when inspired by Brahmagupta, he published his first exposition *Lilavati* (Banerji 1993) that contained his ideas on zero. He mentioned—

> "When a number is added to zero, the result is that number. The square and cube of zero is zero. A number divided by zero has zero for its divisor. When a number is multiplied by zero, the product is zero."

He confirms the properties of 0, laid down by brahmagupta, that $0 + a = a, \frac{a}{0} = \frac{a}{0}, a \times 0 = 0$, where a is any arbitrary number. He further pointed out that—

> In further operations on zero, the multiple of zero, $a \times 0$ in which 0 is a multiplier should be used (just to restore the finite number a acting as the multipli-

[1] Division by zero was featured long ago in ancient Indian mathematics, titled as Zero arithmetic (*sunya ganita*).

> cand of the multiple of 0). However, if there is any operation remaining, zero is merely conceived to be the multiplier, and if zero is also the divisor, the number is considered to be unchanged.

These words were considered to be distinct than those of brahmagupta. As it is believed that bhaskara has used brahmagupta's notion to calculate the motions of the heavenly bodies, and found that it failed sometimes. To fairly distinguish between the two, bhaskara made the following rules to be obeyed while operating zero.

Although $a \times 0 = 0$, the form $a \times 0$ should be used in operations on 0 and if in one of the operations, we encounter with the expression $\frac{a\times 0}{0}$, we should consider $\frac{0}{0} = 1$ so that $\frac{a\times 0}{0} = a$. So, for bhaskara, $\frac{a\times 0}{0} \neq \frac{0}{0}$ but $\frac{a\times 0}{0} = a$. According to him, $a \times 0 = 0$ holds only after the last operation on 0 is performed. He basically introduced the notion of '*impending operations on zero*' in the algebra.

This led Bhaskara to put 0 on the same scale with finite numbers, i.e. for any finite number x, $\frac{a\times x}{x} = a$. He uttered that, '*This method is of great utility in calculating the motions of the heavenly bodies*'. In his zero algebra, he mentioned that "A quantity, divided by zero, becomes a fraction, the denominator of which is zero. This fraction is termed an infinite quantity (*khahara*), i.e. $\frac{1}{0} = \frac{1}{0}$". Resembling with that of infinity, he described that as—

> "There is no change in khahara figure if something is added to or subtracted from the same. It is like, there is no change in infinite Vishnu (almighty) due to dissolution or creation of abounding beings."

This unshaken philosophy of linking the unattained and unbounded infinity with the Lord almighty, has always paved paths to many great philosophical and mathematical Indian minds, who have later on stimulated many more in the same direction.

2.1.2 The Contribution of European Mathematicians

Before the advancements in calculus, people have had no knowledge about the zero and it was considered a mere absence of quantity, specially in Europe. The division by zero came into debate when it was discovered that the foundations of calculus were based on it (discussed in Sect. 3.5). After that, various mathematicians have tried explaining the division by zero as being 'division by absolute nothing or void'. A famous Scottish mathematician and theologian, *John Craig* claimed that 0 cannot be an absolute nothing. His words—

> "Since infinite number of absolute nothings cannot make 1, but by 0 is understood an infinitely small part, as in the calculus, dx is an infinitely small part of x, so that dx is as 0 to x. Not that dx is absolutely nothing, for it is divisible into an infinite number of parts each of which is ddx." (Florian 1919a)

He further believed that there is no quotient (neither 1 or 0 nor unity undivided) arising from the division of 1 by absolute nothing, because there is no division at all. There has always been a point of contention in assigning, zero divided by zero any finite quotient. Some considered this *vanishing fraction as meaningless or undefined probably*, which was put to a question, however no satisfying answer could be encountered. Two famous critics to this idea of vanishing fractions were Rolle and Berkeley, whose analysis conceived such type of fractions to be meaningless. An opposition to Rolle's argument of zero divided by zero, offering a finite quotient, was laid down by his prime antagonist *Varignon*. He exclaimed that—

> "Such fractions as have both their numerator and denominator vanish, or equal to 0, at the same time, may be called vanishing fractions. We are not to conclude that such fractions are equal to nothing, or have no value, for that they have a certain determinate value." (Hutton Charles 2017)

Offering an example of drawing a tangent to certain curves at the point where the two parts cross each other, and as the fractional expression for the sub tangent, by that method, had both its numerator and denominator equal to 0, he supported his argument. Though Rolle discarded it, saying it an absurd expression, and an argument against the method of solution itself[2]. Even Euler defended the opinion of vanishing fractions, by publishing his own work on the differentials in calculus named '*Institutiones Calculi Differentialis*' (in 1755), making an attempt to show that the differentials or infinitesimals dx and dy are zeros and the differential coeffcient $\frac{dy}{dx}$, a vanishing fraction, reducing to a finite variable (Burton 2011). Though he gave a detailed explanation, but mathematicians like J. Gray didn't entertain it, exclaiming the fundamentals of Euler, being dreadfully weak.

In 1828, Ohm accepted that division by zero is unacceptable (Florian 1919b; Romig 1924). He stated that if a is not zero, but b is zero, then the quotient $\frac{a}{b}$ has no meaning, for the quotient multiplied by zero gives only zero and not a. As a consequence, the dogma of this zero division was banished and then onwards the *division by zero is believed to be undefined*. Thereby coming to a conclusion on *the idea of vanishing fractions in Europe* that once posed a problem to even Albert Einstein who exclaimed—"*Blackholes are where God divided by zero. I don't believe in mathematics*" (Gamow 1970).

2.2 The Notion of Symbolism

Undefined is encountered in nearly all the disciplines of mathematics, ranging from geometry to analysis. Following below are some of the popular examples, all of us must have come across, that express the vitality of undefined expressions in mathematics and the need to have a symbol for it too.

[2] This believed mystery was soon explained by John Bernoulli.

Differential equations are often used to represent slopes. Consider the differential equation $\frac{dy}{dx} = \frac{y}{x}$, which assigns a range of different values to the slope, including undefined (vertical lines or no lines) when $x = 0$. Its solution $\frac{y}{x} = m$, is a family of lines passing through origin in the Cartesian co-ordinate that provides different values to the expression $\frac{0}{0}$ depending on the values of the m, as each curve of the family passes through (0, 0).

Consider the differential equation $e^{y''} + y' + \cos x = 0$, of order 2. The highest order derivative in this equation involves an exponential function. Since the expansion of exponential function has infinite terms, so the powers of the second order derivative will keep on altering if more and more terms of the series expansion are included, resulting in an *undefined* degree of the equation.

Euclidean geometry is renowned of its elegance of deriving results just from the existing previous knowledge. In Euclid's '*Elements*', formal definitions are framed using predefined terms and words. But the *Elements* contain a wide use of the terms—*point, line and plane*, which are themselves considered as '*the three undefined terms in geometry*' (Euclid 2002).

Also, Mathematics is evident of new symbols given by the extensive use and application, be it John Wallis (an English mathematician) who is known to give the symbol of *infinity* or Giuseppe Peano who introduced the *intersection sign* (∩), *the union sign* (∪), *the membership sign* (∈), and *existential quantifier* (∃). Some widely known notations of the previous century involve the *set-builder notation* (1960*s*), notation for *inner product*, the *universal quantifier symbol* (∀) (first used by Gerhard Gentzen in 1935) and many more (Florian 1993). Sometimes NaN (Not a Number) is fed in programming languages, incase it encounters any numerical issue, involving a division by zero or an ambiguity. The various examples (many more ahead), wide applications and the history of notations state an appealing need of assigning a symbol/notation to the *undefined*. Also, mathematics has been a language of symbols and adding a symbol for undefined will proliferate the public to get its deep understanding along with its shorthand application and depiction, rather than just writing undefined in the conclusion. Inspired by infinity's lemniscate, we strive forward to provide a symbol "Θ" to denote undefined, and this notation is further employed in the paper, whenever an undefined expression is encountered. This symbol is inspired from the idea of extension of the Real number system, which involves the notion of undefined like $+\infty$ and $-\infty$.

3 Division by Zero

We begin with closely looking at the definitions of undefined, indeterminate and infinity, with some immediate examples to map them with our common understanding.

3.1 Undefined

Mathematically, an expression is said to be undefined, if it has no meaning and therefore can't have an explanation, i.e. '*expressions which are actually not feasible or statements that violate their domain of expression*'. From the functions' point of view, points are often excluded to have a valid 'Domain of Definition' that avoids the ambiguity in the definition of function. Consider

$$f(x) = \frac{1}{\sin x} \quad , \quad \forall x \in \mathbb{R} - \{n\pi, n \in \mathbb{Z}\},$$

whereby the points that make the denominator 0 are removed. Arithmetically, the division by 0, i.e. $\frac{k}{0}$ where $k \in \mathbb{R}$ is one such important example that one encounters (this shall be dealt in detail ahead).

3.2 Indeterminate

An expression or form is referred to be indeterminate if it is *not definitively or precisely determined*. These forms are mostly encountered as the limits of functions in calculus, as the limit definition of the derivative itself is the limit of an indeterminate form. Simply, an expression involving two functions is said to have an indeterminate form, if its limit can not be determined only from the individual limits of the functions. There are basically seven indeterminate limit forms of the functions—

$$\frac{0}{0}, \frac{\infty}{\infty}, 0 \times \infty, \infty - \infty, 0^0, 1^\infty, \infty^0.$$

Since the idea of indeterminate forms is limited to the limit concept of functions only, therefore in arithmetic, all the above mentioned seven as expressions are considered to belong to the family of undefined[3].

3.3 Infinity

Infinity is believed to be a value, actually a symbol, that serves as the bounds for every sequence/number in extended systems like the extended Real line and Complex plane. Usually represented as the ends of the extended systems like the ends of real line or the north pole on the Riemann sphere, there lies two of those $+\infty$ and $-\infty$ in $\mathbb{R}$ and the complex infinity in $\mathbb{C}$. It is often mistook by many as $\frac{1}{0} = \infty$ or $\lim_{x \to 0} \frac{1}{x} = \infty$, this notion is turned down in the paper ahead. Many expressions are assigned this value, like the slowly divergent harmonic series—

[3] Zero to the power of zero (0^0) is an ambiguous case, as its value is not agreed upon by all mathematicians. In algebra, the value is considered to be 1, while in analysis, it is considered as undefined (This paper considers it to be undefined for further use).

$$\sum_{n=1}^{\infty} \frac{1}{n} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots = \infty,$$

just to convey that, there doesn't lie any real number that can serve as an upper bound to this expression.

3.4 Why is $\frac{1}{0}$ Undefined and Not Infinity or Indeterminate

Mathematically, undefined signifies that there is no possible value (or there are infinite possible values), while indeterminate means there is no value that gives the required information. For example, in the equation $ax + by = c$, the relation between x and y is indeterminate because one can't determine it without knowing the values of a, b and c.

Consider the following condition, in order to determine the value of an expression of type $\frac{1}{0}$ in arithmetic. Let us safely assume (if possible)

$$\frac{n}{0} = k, \forall n \in \mathbb{Z}, k \in \mathbb{R}.$$

Case 1: If $n \neq 0$, multiply both the sides by 0, we get $n = 0$ (impending operation) which is a contradiction, and therefore implies the undefined nature of $\frac{1}{0}$.
Case 2: If $n = 0$, multiply both the sides by 0, we get $0 = 0$ (impending operation) and this is true for all $k \in \mathbb{R}$. So we are not able to assign any particular value to the expression $\frac{0}{0}$, and hence referred as undefined in arithmetic.

In simple words, for the real valued function

$$f(x) = \frac{1}{x},$$

the domain of definition is $\mathbb{R} \setminus \{0\}$. This is quite similar to why the Group of non-zero real numbers with multiplication as group operation viz. $(\mathbb{R} \setminus \{0\},*)$ is also defined over the same, due to inability to have an inverse for 0. As multiplication is taken to be an inverse process of division, and $x \times 0 = 0$, $\forall x \in \mathbb{R}$, therefore division by 0 breaks the uniformity of assigning unique values while dividing (that's why we used impending operations in above cases). So the usually encountered expression $\frac{1}{0} = \infty$.

In a similar fashion, with the idea of limit of a function, the expression can also be written as—

$$f(x) = \lim_{x \to 0} \frac{1}{x}.$$

For the prediction of answer, it is as strange as before. But, because of the limiting nature of this real valued function, the division by 0 results in two notions that are

poles apart, i.e. the two infinities. Poles apart in the context that, approaching the limit from left side on the Real line gives us $-\infty$ that serves as the lim inf[4] of range set for all the unbounded below functions or sequences, and approaching from right gives us $+\infty$, i.e. the lim sup[5] for the unbounded above's.

3.5 Close to Indeterminate Form

Division by zero has been of significant value to us. Let's assume a line that intersects the graph of an arbitrary function $F(x)$ at two points x_1 and $x_2 = x_1 + h$. Then by the definition, the derivative is the slope of the tangent line which is given as—

$$\frac{d}{dx}F(x) = \lim_{h\to 0}\frac{h}{F(x_2) - F(x_1)} = \frac{0}{0}.$$

The core essence of derivative relies on the limiting nature of $\frac{0}{0}$ or the division by 0 idea, so without this conception it would have been really difficult for us to think of a calculus devoid of derivatives.

Broadly from the *perspective of Arithmetic*, we may consider division by zero to be a subclass of indeterminate expressions, as the idea of indeterminate is restricted to that of functions only. However from the *perspective of Functions*, the foremost difference is that, no matter how you try to manipulate an undefined expression, it is always going to be as weird as before. While there are instances where indeterminate functions can be manipulated in a way to obtain meaningful results, when the limit approaches $\frac{0}{0}$. For example: $\lim_{x\to 0}\frac{\sin x}{x} = 1$.

The symbol ∞ is not a number (Bartle and Sherbert 2011), rather an action of an unceasingly growing function. So encountering a function that tends to infinity, reflects a never ending positively growing curve. A meticulous glance would give an idea that all the other 6 indeterminate forms may be manipulated to get $\frac{0}{0}$ form. Like some functions of the form $f(x)^{h(x)}$ with limit as 1^{∞}, can be manipulated to go to a particular number (depending on the nature of the two functions) using L'Hôpital rule as demonstrated below—

$$\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e^{\lim_{x\to\infty}\frac{\log\left(1+\frac{1}{x}\right)}{\frac{1}{x}}} = e^{\lim_{x\to\infty}\frac{\frac{-1}{x^2}}{\left(1+\frac{1}{x}\right)\times\frac{-1}{x^2}}}$$

$$\Rightarrow \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e.$$

This is all due to the limiting nature of the functions that $\frac{0}{0}$ as a fraction is undefined, while the form of the limit is indeterminate and therefore can be manipulated.

[4] lim inf of a sequence $<x_n>$ is defined as—$\liminf x_n := sup\{inf\{x_m : m \geq n\} : n \geq 0\}$.

[5] lim sup of a sequence $<x_n>$ is defined as—$\limsup x_n := inf\{sup\{x_m : m \geq n\} : n \geq 0\}$.

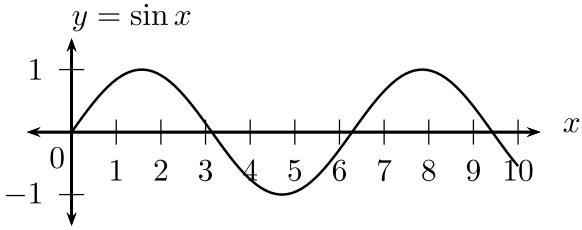


Fig. 1 The graph shows unpredictable nature of sin x as $x \to \infty$

4 Unlike Nature in Number Theory and Functional Analysis

This section analyses the classes in the reference of number theory and functional analysis, i.e. the vivid idea in Arithmetic and as Functions. Let's recall the definitions, while emphasizing at the tables below.

In the light of Table (a), Infinity is a notion representing the largest number, with its existence and determination assured (because it can be put at the end of the extended Real Line or on the north pole of the Riemann Sphere). Due to the non-existence idea of indeterminate in arithmetic, its existence and determination makes no sense. Having no hint of existence, undefined remains quite troublesome to deal with.

(a) From the perspective of Number Theory

	Existence	Determination
∞	Yes	Yes
Indeterminate	No	No
Undefined	Not Sure	No

In the light of Table (b), considering infinity as just a notion for the limit of an unceasingly growing real valued function, its existence is believed to prevail but no assured determination can be claimed. In the limelight for the limiting nature of functions, indeterminate expressions exists and can be manipulated to give a precise answer only if the limit prevails. Whereas undefined expressions in functions, expresses the ambiguity that arises while finding values of the function at various points.

(b) From the perspective of Functional Analysis

	Existence	Determination
∞	Yes	No
Indeterminate	Yes	Yes
Undefined	Not Sure	No

In the view of Number Theory, *determination* implies *existence*, and therefore *no determination* implies *nothing can be said*. So we may consider that '*Arithmetically, the class of indeterminate is contained in that of undefined*', as every indeterminate form is undefined in the sense that neither of them can be put at the real line. So

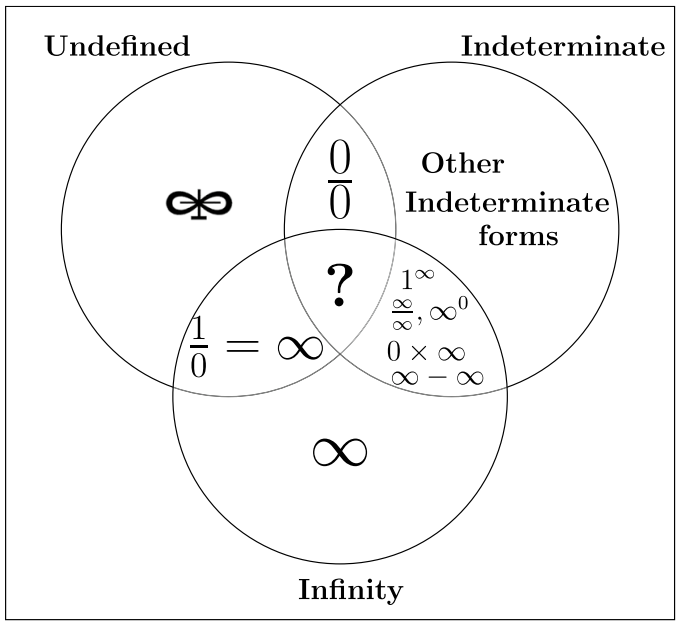


Fig. 2 The Relation between Undefined, Indeterminate and Infinity

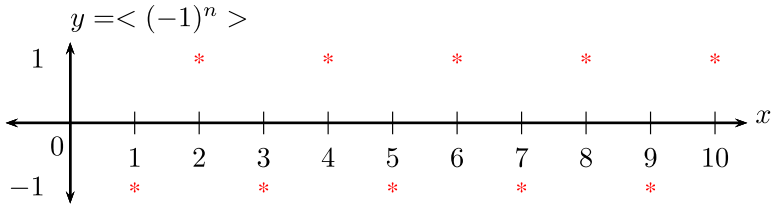


Fig. 3 The oscillating nature of the sequence $<(-1)^n>$

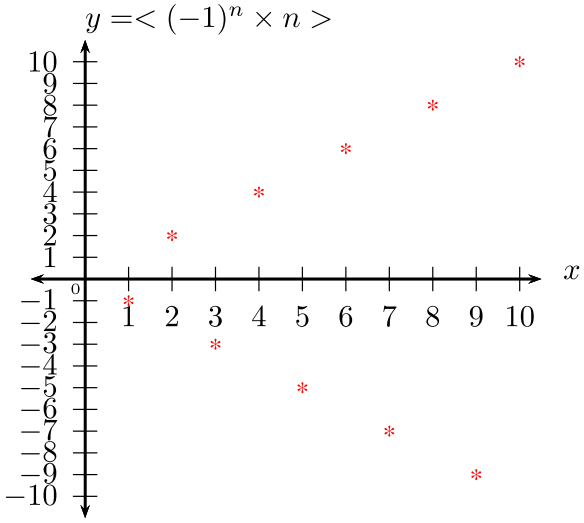


Fig. 4 The infinite oscillations of the sequence $<(-1)^n \times n>$

as far as the domain of number theory is concerned, we can discard the opinion of indeterminate at hand, exclaiming the seven indeterminate forms of limit to be undefined only.

But this is not the case with the functions because the limiting nature of functions is different than just evaluating the functions with their individual limits. Example: $\lim_{x\to\infty} \sin x$ *is* ∞ *and not indeterminate* (refer to Fig. 1), while an expression like 1^0 (an indeterminate form) is undefined in the light of number theory.

Also, undefined algebraic expressions don't correspond to any indeterminate form in both arithmetic and functions. Like $\frac{1}{0}$ is undefined as a fraction and it does not correspond to any indeterminate form in functions, because any limit that gives rise to this form will diverge to infinity. This is due to the fact that limit of a function is always determinate, i.e. it either exists or not, It is only the form of the limit that is indeterminate and therefore can be manipulated from one form to another. The following self explanatory Venn Diagram (Fig. 2), shows a more close interwoven relation among the three classes, linking the concepts established above.

Following below are some interesting and noteworthy results along with examples for the sequences, in the analysis of real valued functions to make things clearer.

Axiom *For any finitely or infinitely oscillating sequence* $<a_n>$, $\lim_{n\to\infty} a_n$ *is undefined.*

Example 1 $<a_n> = <(-1)^n>$ is a finitely oscillating sequence (refer to Fig. 3), and clearly $\lim_{n\to\infty} a_n = \not\exists$.

Example 2 $<b_n> = <(-1)^n \times n>$ is an infinitely oscillating sequence (refer to Fig. 4), and $\lim_{n\to\infty} b_n = \not\exists$.

Axiom *Given a divergent Sequence* $<a_n>$, $\lim_{n\to\infty} a_n$ *may or may not be undefined.*

Example 3 $<a_n> = <n>$ is a divergent sequence, and $\lim_{n\to\infty} a_n = +\infty$ which belongs to the extended real line.

Example 4

$$<b_n> = \begin{cases} n \text{ if n is prime} \\ p \text{ otherwise, where } p \text{ is the greatest prime divisor of } n \end{cases}$$

is a divergent sequence, but $\lim_{n\to\infty} b_n = \not\exists$.

The various examples illustrated above give us a fair enough idea about the nature of various sequences and how the concept of undefined often confronts in sequences too, like infinity. It is just a mere absence of notation for undefined that the *undefined tails of various sequences* are not put much to light.

5 Conclusion

In this paper, we have examined that in the light of number theory, indeterminate is a subclass of undefined or there is no idea of indeterminate at all. While for functions, the notions of undefined and indeterminate are distinct. We have also discussed about the behaviour of tail of oscillatory and divergent sequences. Discussing

the utility and numerous encounters of undefined in daily life mathematics, we have begot a notation $\frac{1}{0}$ for the same in this work.

We aspire the work to open the realm of new branches, in the field of division by zero. All those who feel like sharing their ideas and reviews about the same are hearty welcomed.

References

Banerji HC (1993) Bhaskaracharya's lilavati with Colebrookes translation and notes, 2nd edn. Asian Educational Series, New Delhi

Bartle Robert G, Sherbert Donald R (2011) Introduction to real analysis, 4th edn. John Wiley & Sons, New York

Brahmagupta (1966) Brahmasputa Siddhanta. In: Acharyavara RSS (ed) Indian Insitite of Astronomical and Sanskrit Research, New Delhi

Burton David M. (2011) The history of mathematics: an introduction, 7th edn. McGraw-Hill, New York

Cajori F (1919) A history of the conceptions of limits and fluxions in Great Britain, from Newton to Woodhouse. Chicago Open Court Pub Co, Toronto

Cajori F (1919) A history of mathematics, 2nd edn. The Macmillan Company, New York

Cajori F (1993) A history of mathematical notations. Dover Publications, New York

Euclid (2002) Euclid's elements: all thirteen books complete in one volume. Green Lion Press, Santa Fe

Gamow G (1970) My world line : an informal autobiography, 1st edn. Viking Press, New York

Hutton C (2017) A philosophical and mathematical dictionary: containing an explanation of the terms, and an account..., vol 1. Forgotten books, London

Romig HG (1924) Discussions: early history of division by zero. Am Math Monthly 31(8):387–389

Ufuoma O (2017) On the operation of division by zero in Bhaskara's framework: survey, criticisms, modifications and justifications. Asian Res J Math

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

Proceedings of the 6th International Conference on Communication and Electronics Systems (ICCES-2021)
IEEE Xplore Part Number: CFP21AWO-ART; ISBN: 978-0-7381-1405-7

DESIGN AND ANALYSIS OF RING MULTIBAND METAMATERIAL ABSORBER

Priyanka Jain
Department of Electronics and Communication,
Delhi Technological University, New Delhi, India-42
priyankajain@dtu.ac.in

Keshav Bhati
Department of Electronics and Communication,
Delhi Technological University, New Delhi, India-42
keshavbhati_2k19moc01@dtu.ac.in

2021 6th International Conference on Communication and Electronics Systems (ICCES) | 978-1-6654-3587-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICCES51350.2021.9489035

***Abstract*–In this paper a metamaterial absorber (MA) inspired from a split-ring structure has been designed that exhibits three resonant peaks in the C-band and starting of X-band range. Proposed structure for the absorber is a three-layer rectangular split-ring with different absorption peaks at 8.488, 8.02, 7.51 GHz with 99.25%, 95.11%, 93.93% absorptivity respectively. MA is designed on the low-cost FR-4 material with dielectric constant 4.4 and substrate thickness 0.5 mm. Performance of the absorber is evaluated in terms of absorptivity related to three layers, material parameters and normalized impedance.**

Keywords – Absorptivity, Split-Ring, Unit-Cell Metamaterial Absorber (MA), Metallic loops, Slot.

I. INTRODUCTION

In the last few years, metamaterial have become an important area of research for their applications in the field of microwave filters [1-2], super lenses [3], antenna design [4-7], cloaks [8]and as an electromagnetic absorber [9-17] etc. Metamaterials consists of multiple unit cell whose dimensions are much smaller than λ/4 [10]. Conventional electromagnetic absorbers such as Jaumann and Salisbury screen absorber are very bulky in nature but stealth technology and radar systems demand absorbers to be compact and light weight, that's where metamaterial absorbers offer an advantage due to their ultrathin thickness with near perfect absorption [11].It also found application in various areas like reduction of radar cross section (RCS) antennas [4-7], radome[12], EMI/EMC application[13] etc.MA is a such metallic structure which can restrict the reflection and transmission of the electromagnetic (EM) wave within a certain range of frequencies. They are generally made in a n x n array of unit cells. Dimension of these unit cells can be taken based on their operating frequency.

Rest of the paper is organized as follows: Section II deals with literature review and basic theory related to metamaterial absorber is discussed in section III. Design of the proposed absorber and its simulation results are being explained in section IV. Section V highlights the parametric analysis and finally conclusions of the work are presented in section VI.

II. LITERATURE OF METAMATERIAL ABSORBERS

The first metamaterial absorber was introduced by Landy *et. al* in 2008[9], since then narrowband [14], multiband [15-16] and broadband absorbers [17-19][23] are the hot topic for researchers.

Huang and Chen [20] proposed a triple-band MA that uses three square rings having different sizes distributed in two substrates based on multi-layer. A three-layer absorber has been proposed in [4] that has absorptivity greater than 90%. In another work a flower shape metamaterial absorber is used to increase the isolation between the MIMO antennas [21]. In [22] metamaterial absorber has been proposed for radar cross section (RCS) reduction as well as improving the isolation between elements MIMO antenna for Radiolocation and Aeronautical Radio Navigation Application

A multi-layer resonator with different shapes/ same shape but with different sizes offers a near perfect multiband/ broadband response. Multiple resonators in unit cell generated multiple discrete frequency points. Continuous and wide absorption spectrum can be achieved by multiple frequency points with narrow frequency interval [23]. The objective of this paper is to propose a multi-layer split-ring type MA unit cell which can have multiband response and to study the effect of varying split size on absorptivity.

III. BASIC THEORY OF ABSORBER

As we know that whenever an electromagnetic wave strikes a surface it may get transmitted, reflected, scattered or absorbed. To absorb the electromagnetic radiation, a metamaterial absorbs, controls the loss components of the metamaterial, relative permittivity (ε_r) and permeability (μ_r). The absorbance (A) of a surface can be given as:

$$A = 1 - R - T \quad (1)$$

Here,
Ris Reflectance which is obtained from $|S_{11}|^2$
T is Transmittance obtained from $|S_{21}|^2$

Since in the proposed structure the bottom of the substrate is completely coated with the copper, it further restricts the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:08:05 UTC from IEEE Xplore. Restrictions apply.

propagation of wave, which indicated that transmittance i.e., $|S_{21}|^2 = 0$. Therefore, equation (1) can be rewritten as:

$$A = 1 - |S_{11}|^2 \qquad (2)$$

IV. DESIGN AND SIMULATION

Suggested MA unit cell is a symmetrical structure composed of three metallic loops and a ground plane. The unit cell is designed on a low cost FR4 substrate which has a relative permittivity of 4.4. All the three metallic strips are separated by a dielectric with a thickness of 0.5 mm making the structure with total thickness of 1.5 mm. The metal used is copper with a conductivity of 5.8×10^7 S/m. Floquet Port and periodic boundary conditions are used to simulate the MA structure in Ansys HFSS. Fig. 1 shows a basic three-layer MA structure. The simulated result of reflection coefficient(S_{11}) is shown in Fig. 2. It is observed that there is a sharp decline at around 5GHz with a minimum value of S_{11} is -40.11 dB at 5.04 GHz.

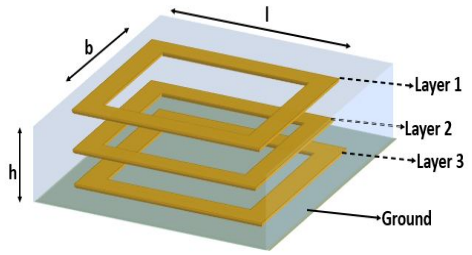


Fig. 1. Geometry of three-layer MA unit Cell

As depicted in Fig. 3, the structure has 99.99% absorptivity at 5.04 GHz. It is observed that there is another minor absorption peak at around 5.4 GHz with 95% absorptivity and absorption starts to decline on the higher frequencies. It can be interpreted from the Fig. 3 that MA structure is perfect absorber at 5.04 GHz and shows absorptivity more than 90% at 5.4Ghz.

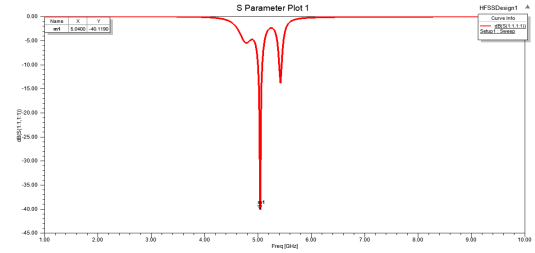


Fig. 2. Plot of S_{11} w.r.to frequency (GHz)

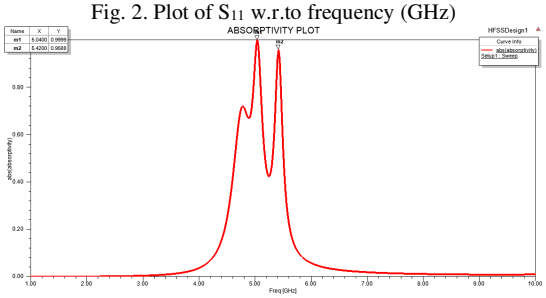


Fig.3. Absorptivity result of MA unit cell

To design a split-ring MA, we have introduced a slot in the middle one side of metallic ring of each layer of the MA. Thus, the final structure will be similar to a split ring structure.

In the first stage, a 1mm wide slot has been introduced in the middle of one side on the top layer of the ring structure, corresponding split ring structure is shown in Fig. 4.

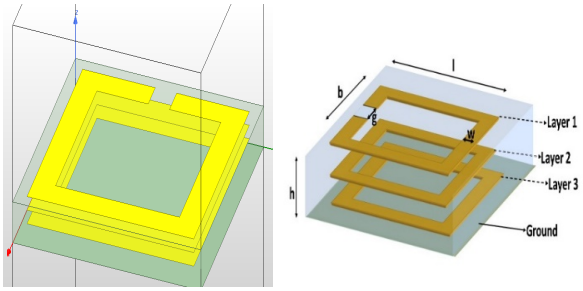


Fig.4. Geometry of a 3-layer MA with a slot on top layer

A plot of reflection coefficient S_{11}has been shown in the Fig.5. It is observed from the Fig. 5 that the reflection coefficient is less than -10 dB at around 5.03 GHz and 8.48 GHz and with a minimum value of -20.7 dB at 5.03 GHz which is the operational frequency of this structure. It is observed from the Fig. 6 that in comparison with the original structure the absorptivity plot now shows multiple peaks with two significant peaks with 99.16 % and 94.64 % at 5.03 and 8.488 GHz respectively. The results shows that due to multiple layer structure multi frequency resonance is obtained in MA.

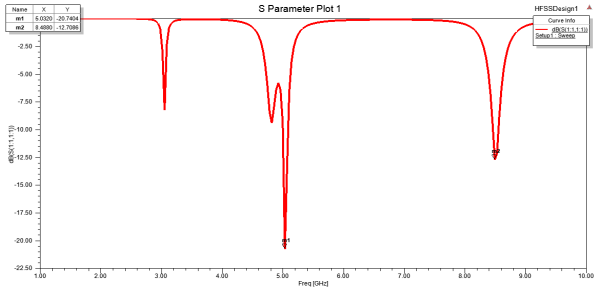


Fig. 5. Plot of S_{11} vs frequency of MA unit cell with one slot

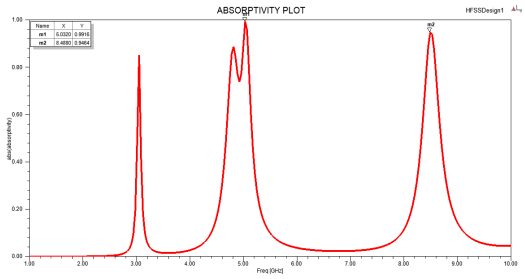


Fig. 6. Absorption result of the MA unit cell with a slot

In the second stage, in addition with the first slot in the top layer, another slot has been introduced in the middle layer.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:08:05 UTC from IEEE Xplore. Restrictions apply.

The slot in the middle has same dimension as on the top layer and it is exact below the top layer slot. As depicted in Fig.7, the structure now consists of slots on the top layer as well as on the middle layer. Absorptivity curve is shown in Fig. 8. With the two slots in the layers, resonance frequency shift has been observed and corresponding three significant absorption peaks at 3.19, 4.81, 8.45 GHz with 99.93 %, 98.26 %, 96.67 % absorptivity respectively has been achieved. It is also observed form the fig. 8 that absorptivity has been increased with the increase in the slots.

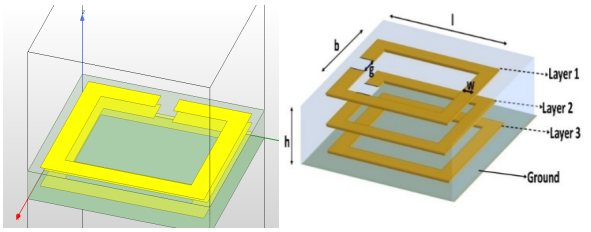


Fig. 7. Geometry of MA unit cell with 2 slots in layers

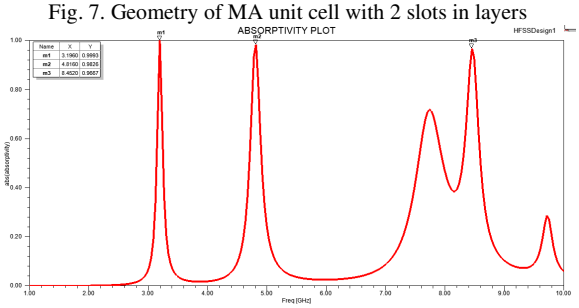
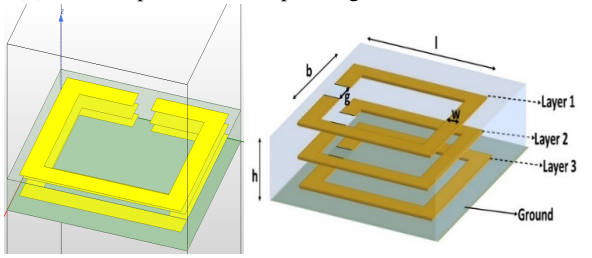
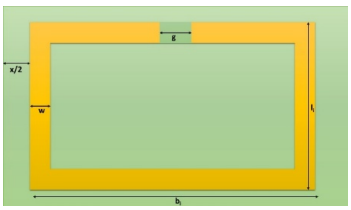


Fig. 8. Absorption plot of unit cell with slots in the middle and top layer.

In third stage, a three-layer split ring MA structure have made by introducing another 1mm wide slot in the bottom layer so as to make all three layers as a split ring. A schematic diagram of three-layer split ring structure is shown in Fig. 9(a). Fig. 9(b) shows top view of the split-ring structure.



(a)



(b)

Fig. 9. Geometry of (a) three-layer split ring MA unit cell (b) Top view of the unit cell

As depicted in Fig. 10, the split ring MA unit cell shows different absorption peaks at 8.488, 8.02, 7.51 GHz with 99.25%, 95.11%, 93.93% absorptivity respectively. All the major frequency points are at very close interval. It can be observed that MA unit cell have absorptivity greater than 90% for multiple frequency. It shows absorptivity greater than 99% at 8.48 GHz means that electric and magnetic are strongly in resonance(ε=μ) at around 8.48 GHz and near-perfect impedance-matched to the free space, resulting in near-zero reflection as shown in Fig. 11 [24].

Fig. 11 shows the reflection coefficient(S_{11}) and absorptivity curves. It can be observed from the Fig. 11 that the absorptivity approaches near unity when the reflection approaches near zero.

It can also be observed from the Fig. 10 that the structure is not showing a broadband absorption instead showing multiple absorption peaks at different frequencies lying in two band C-band and X-band therefore, the structure is a multiband structure.

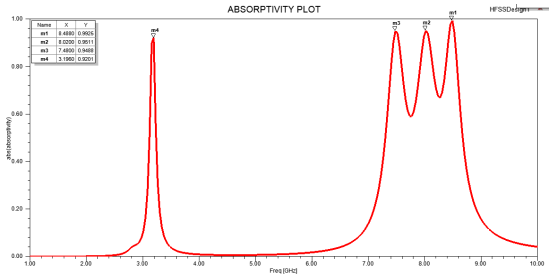


Fig. 10. Absorptivity result of split ring MA unit cell.

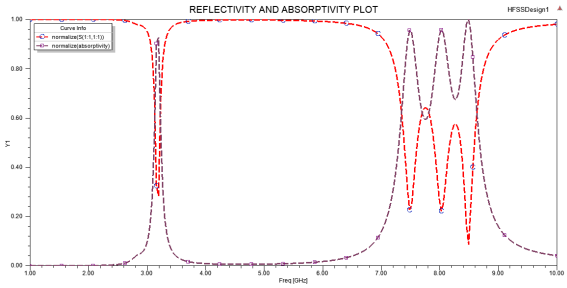


Fig. 11. Simulated results of reflectivity and absorptivity of split-ring unit cell.

Finally design parameters of the unit cell are as follows:

TABLE I. Design Parameters of the unit cell

Parameter	Value	Description
h	1.5 mm	Substrate Height (three layers)
b	8.8 mm	Metal loop width
w	1 mm	Metal loop size
x	1 mm	Distance along b
g	1 mm	Slot gap size

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:08:05 UTC from IEEE Xplore. Restrictions apply.

TABLE II. Length and size of metal loops

	l_i	w_i
LAYER - 1	7.8 mm	1 mm
LAYER - 2	7.2 mm	1.2 mm
LAYER - 3	7.8 mm	1 mm

A comparison of all the structure's absorptivity plots is shown in Fig. 12. It can be observed that final three-layer split-ring structure shows multiband absorption characteristics with three distinct absorption peaks having absorptivity greater than 90%. The first structure without any slot shows only one absorption peak and after introduction of the slots MA starts showing multiband absorption. The operating frequency is also shifted from 5.04 GHz to 8.48 GHz in the final structure.

TABLE III. Absorptivity results with different no. of slots

No. of slots	Frequency (GHz)	Absorptivity
0	5.04	99.99 %
1	5.03	99.16 %
	8.48	94.64 %
2	3.19	99.93 %
	4.81	98.26 %
	8.45	96.67 %
3	7.51	93.93%
	8.02	95.11%
	8.48	99.25 %

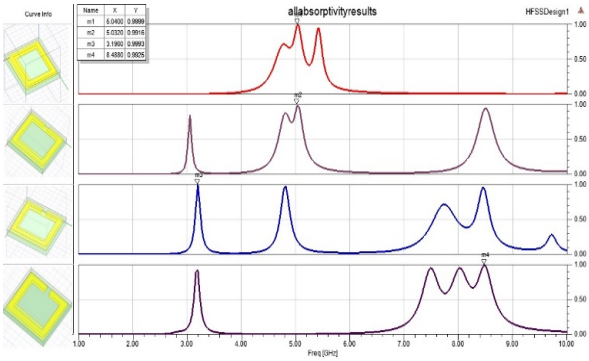


Fig. 12. Comparison of absorptivity of all structures.

V. PARAMETRIC ANALYSIS

Further the effect of varying slot width(g) on absorptivity of the structure is studied in this section. Multiple values of slot width(g) have been taken such as 0.5, 0.75, 1, 1.25, 1.5 mm. The simulated results variation of parameter 'g' on absorptivity is shown in the Fig.13. The Multiband absorption characteristics remain same but with a slight change in operational frequency having maximum absorption. Varying the slot width changes the structure's capacitance and inductance when considering the structure as RLC circuit which therefore leads to the change in operating frequency.

TABLE IV. Absorptivity results with different slot size

Slit width (G)(mm)	Absorptivity	Frequency (GHz)
0.5	99.43 %	7.73
0.75	99.14 %	8.36
1	99.25 %	8.48
1.25	97.98 %	8.48
1.5	99.85 %	7.66

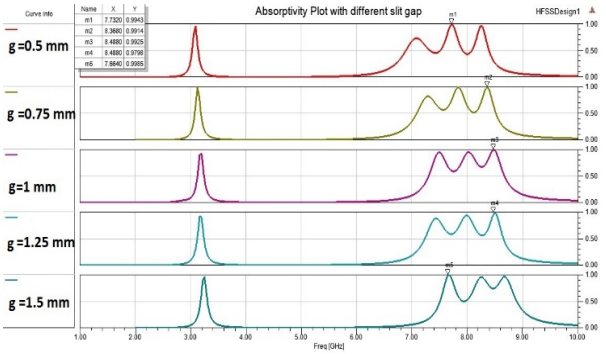


Fig. 13. Absorptivity variation with variation of slot size

VI. CONCLUSION

In this paper, a three-layer split-ring metamaterial absorber (MA) unit cell has been proposed. The split-ring MA unit cell is being derived from a conventional three-layer ring MA structure by introducing slots in the metallic loops of the structure. The three-layer structure is a signal resonance frequency structure with 99% absorptivity at 5.04 GHz. It is observed from the simulated results that by introducing slots in the MA cell, it becomes a multiband electromagnetic absorber. The split-ring MA cell shows multiband absorption with absorption peaks at 8.488, 8.02, 7.51 GHz with 99.25%, 95.11%, 93.93% absorptivity respectively. Further, the effect of varying slot size on the absorptivity has also been studied and it has been observed that by changing the slot width only shift in the frequency can be obtained without any significant change in the absorptivity.

REFERENCES

1. M. Gil, J. Bonache, F. Martín, "*Metamaterial filters: A review*", *Metamaterials*, Vol. 2, Issue 4, pp. 186-197, 2008.
2. Statsenko, Lyubov G., et al. "*Designing Microwave Filters Using Metamaterials*." Key Engineering Materials, vol. 806, Trans Tech Publications, Ltd., June 2019, pp. 167–172.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:08:05 UTC from IEEE Xplore. Restrictions apply.

3. J. B. Pendry, "*Negative Refraction Makes a Perfect Lens*," Phys. Rev. Lett., vol. 85, no. 18, pp. 3966–3969, Oct. 2000.
4. Junyi Ren, Shuxi Gong, "*Low-RCS Monopolar Patch Antenna Based on a Dual-Ring Metamaterial Absorber"* IEEE Antennas and Wireless Propagation Letter*s,* vol. 17, no. 1, January 2018.
5. S. Genovesi, F. Costa, and A. Monorchio, "*Low-profile array with reduced radar cross section by using hybrid frequency selective surfaces*," IEEE Trans. Antennas Propag*.,* vol. 60, no. 5, pp. 2327–2335, May 2012.
6. J. Zhang, J.Wang, M. Chen, and Z. Zhang, "*RCS reduction of patch array antenna by electromagnetic band-gap structure*," IEEE Antennas Wireless Propagation Letter*s* vol. 11, pp. 1048–1051, 2012.
7. S. J. Li, J. Gao, X. Y. Cao, Y. Zhao, Z. Zhang, and H. X. Liu, "*Loading metamaterial perfect absorber method for in-band radar cross section reduction based on the surface current distribution of array antennas*," Microwave Antennas Propagation, vol. 9, no. 5, pp. 399–406, Apr. 13, 2015.
8. D. Schurig, J. J. Mock, B. J. Justice, S. A. Cummer, J. B. Pendry, A. F. Starr, and D. R. Smith, "*Metamaterial electromagnetic cloak at microwave frequencies*," Science, vol. 314, pp. 977–980, 2006.
9. N. I. Landy, S. Sajuyigbe, J. J. Mock, D. R Smith, W. J and Padilla, "*Perfect metamaterial absorber,*" Phys. rev. lett., vol. 100, issue 20, 204702, 2008.
10. R. S. Kshetrimayum, "*A Brief intro to Metamaterials*," IEEE Potentials, vol. 23, no. 5, pp. 44-46, Dec. 2004-Jan. 2005.
11. Priyanka Garg and Priyanka Jain, "*Analysis of a novel Metamaterial Absorber using equivalent circuit model operating at 3.5 GHz*", 2020 IEEE International Conference on Computational Electromagnetics (ICCEM 2020), pp. 246-247, 2020.
12. Öziş, E. & Osipov, Andrey &Eibert, Thomas. (2017). "*Metamaterials for Microwave Radomes and the Concept of a Metaradome: Review of the Literature*." International Journal of Antennas and Propagation. 2017. 1-13. 10.1155/2017/1356108.
13. P. Munaga, S. Ghosh, S. Bhattacharyya, D. Chaurasiya, and K. V.Srivastava, "*An Ultra-thin Dual-Band Polarization-Independent Metamaterial Absorber for EMI/EMC Applications*," 9th European Conf. Antenn. Prop. (EuCAP), Lisbon, 2015, pp. 1–4.
14. Min Zhong, "*Design and measurement of a narrow band metamaterial absorber in terahertz range,*" Optical Materials, Volume 100,2020,109712, ISSN 0925-3467,
15. Xiebin Yu, Yaoliang Song, and Shicheng Fan, "*Research on a multiband metamaterial absorber*", AIP Conference Proceedings 1890, 040102 (2017).
16. R.M.H. Bilal, M.A. Baqir, P.K. Choudhury, M.M. Ali, A.A. Rahim, W. Kamal, "*Polarization-insensitive multi-band metamaterial absorber operating in the 5G spectrum*", Optik, Volume 216,2020,164958, ISSN 0030-4026.
17. Y. Xu, B. Zhang, J. Duan and Y. Tian, "*Design of Broadband Metamaterial Absorber for C Band and X Band*," 2018 IEEE International Conference on Computer and Communication Engineering Technology (CCET), 2018, pp. 159-163.
18. Zilong Zhang, Lei Zhang, Xiqiao Chen, Zhuang Wu, Yaoyi He, YangyangLv, Yanhong Zou, "*Broadband metamaterial absorber for low-frequency microwave absorption in the S-band and C-band*", Journal of Magnetism and Magnetic Materials, Volume 497,2020,166075, ISSN. 0304-8853.
19. Parul Dawar, N.S. Raghava and Asok De, "*UWB Metamaterial-Loaded Antenna for C-Band Applications",* Hindawi International Journal of Antennas and Propagation, Vol. 2019, Article ID 6087039, 13 pages.
20. Liang Huang and Hongsheng Chen, "*Multi-Band and Polarization Insensitive Metamaterial Absorber*," Progress in Electromagnetics Research, Vol. 113, 103-110, 2011.
21. Priyanka Garg and Priyanka Jain, "*Isolation Improvement of MIMO Antenna Using a Novel Flower Shaped Metamaterial Absorber at 5.5 GHz WiMAX Band*", IEEE Transactions on Circuits and Systems—ii: express briefs, vol. 67, no. 4, april 2020
22. Saxena, Gaurav & Awasthi, Yogendra & Jain, Priyanka. (2021). "*Design of metasurface absorber for low RCS and high isolation MIMO antenna for radio location & navigation*". AEU - International Journal of Electronics and Communications. 133. 153680. 10.1016/j.aeue.2021.153680.
23. Wen, Ding-e &Helin, Yang & Ye, Qiwei& Li, Minhua& Guo, Linyan& Zhang, Jianfeng. (2013). "*Broadband metamaterial absorber based on a multi-layer structure*". Physica Scripta. 88. 015402.
24. Cheng, Y., Yang, H., Cheng, Z. *et al*. "*Perfect metamaterial absorber based on a split-ring-cross resonator*". Appl. Phys. A 102, 99–103 (2011).

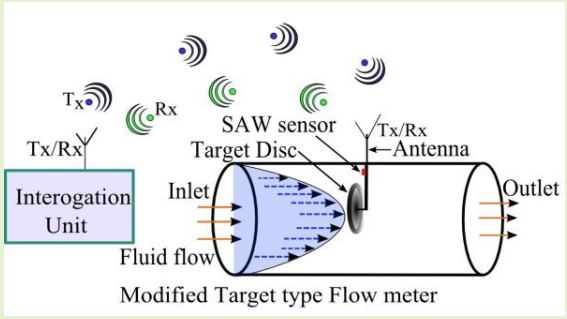
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:08:05 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2021.3104863, IEEE Sensors Journal

Design and Development of SAW Sensor based Wireless Target type Flow Transmitter

Praveen Maurya, *Member, IEEE*, Sayyed Faizan Ali, *Member, IEEE* and *Nirupama Mandal*, Senior *Member, IEEE*

***Abstract*— Target flowmeter is commonly used for measuring flow rates, but there are some shortcomings of the conventional target flowmeter, such as complex structure, low sensitivity, non-linearity. Here, a Surface Acoustic Wave (SAW) sensor-based target flowmeter has been proposed where a SAW sensor is placed on the target rod of the flowmeter. When liquid flows through the pipe in contact with the target rod, the force is produced. This force then acts on the target rod and deflects the target rod from its original position, which results strain in it. SAW sensor senses the strain level at different flow rates, and the output is in terms of delayed time response of the order of a nanosecond. The output is non-linear with respect to the flow rate and it is linearized with the help of a neural network technique. This minimizes the non-linearity error up to 0.27% of full-scale output (FSO). The modification in the target type flowmeter is done with the help of the SAW sensor in such a way that the flow transmitter can transmit sensed information through the wireless channel. Simulation and mathematical approaches are reported in this paper for the range of 0 to 1200 LPH.**



***Index Terms*— Flow measurement, target type flowmeter, SAW sensor, time response.**

I. Introduction

In oil, gas, chemical and process industries flow measurement is an important parameter to control the process. In industries, flow measurement is done by various traditional flowmeters like electromagnetic, ultrasonic, differential pressure flowmeter, rotameter, and fiber optic-based flow meter etc. Electromagnetic flowmeter works on the principle of the Faraday law of electromagnetic induction. It is used for the measurement of conductive liquid [1, 2]. The limitation of this instrument is that it is affected by the external magnetic field and is not capable of measuring the flow rate of the low conductive liquid. Similarly, the ultrasonic flowmeter is based on the measurement of sound velocity. The velocity of sound is influenced by the temperature, which generates a partial error in flow measurement. A differential pressure flowmeter is used for liquid, gas and steam. There are no moving parts, and it is used for a wide range of sizes. But it supports a short span due to the square root relationship between flow rate and differential pressure. Rotameter is a simple, cost-effective, and provides linear scale, but when opaque fluid is used, the float may not be visible, must be installed in the vertical position and is not well suited for pulsating services. Fiber optic-based flow meter is small in size, passive, low attenuation, immunity to electromagnetic interference, but it is very costly, difficult to install, fragile and needs a complex detection system. In the process industry, there are a number of different types of conventional instruments used to measure the flow rate but they have limitations. To overcome those limitations and to design sophisticated flow sensors to fulfil the requirement of modern process plants different research works are being carried out by different researchers. Sunita Sinha et al. [3] have developed a PC based flow measurement system by modifying conventional rotameter. Modification is done in such a way the rotameter is capable to transmit the flow rate information in terms of voltage to the remote location, and it can be monitored in PC at the control room.

To control the process variable from remote locations, the measured information should be transmitted from a process location to a remote location. Nowadays, data transmission is possible in three ways wired, wireless and optical. Every transmission method has some limitations depending upon the complexity and cost. At the same time, for harsh and inflammable industries, special precautions are to be taken to design a flowmeter. In such an environment, an SAW-based flowmeter can work satisfactorily. In this paper, we have proposed wireless passive flow transducers by modifying the target type flow measuring instrument with the SAW sensor. SAW sensor is a device that does not require any external power supply for operation. The sensor is very robust, small (of the order of mm) and lightweight. It can work in a very harsh environment and at different temperatures level. SAW sensor is powered by an external RF wave [4], [5]. Many researchers have proposed various applications using SAW sensors. S G Joshi et al. [6] have used the SAW oscillator-based device to measure the rate of flowing fluid using the delay line principle. A SAW device using ultrasonic Lamb

1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:18:54 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2021.3104863, IEEE Sensors Journal

waves Wolf-Eckhart Bulst et al. [7] have discussed different sensing techniques based on the SAW sensor, SAW transponder and analyze the effect of temperature on the SAW sensor. A physical quantity like pressure can also be measured by using SAW sensor. Most of the SAW sensors are used for high-pressure measurement. Taleb Moazzeni et al. [8] have developed a system that can measure the flow rate of liquid in an extremely harsh environment. The system was developed by using the SAW devices. Real-time measurement of process parameters using SAW sensor has been proposed by R. Stoney et al. [9]. This passive sensor can work effectively wirelessly in harsh environments. Yizhong Wang et al. [10] have performed the theoretical analysis on the SAW sensor for flow measurement. The flow rate is extracted from a pipe by using a change in the pressure of the SAW sensor. V. Kalinin [11-13] have proposed and designed a number of SAW-based for different applications. This researcher has proposed a SAW-based strain sensor [11] and predicted the shear strain using simulation. The stability of SAW sensor, especially the stability of frequency output, is estimated and the drift of the sensor per year is also estimated by V. Kalinin et al. [12]. V. Kalinin [13] has also proposed another SAW sensor-based technique for multivariable measurement. This passive resonant sensor can measure strain, as well as temperature, torque and pressure. M. Rodriguez et al. [14] have presented an orthogonal frequency coadding (OFC) passive wireless sensing system consists of a transceiver that sends a wireless interrogation signal to the SAW sensor. The performance of OFC based SAW is a function of the signal to noise ratio as derived from the analog sensor. The extracted temperature has a precision better than 0.001◦C. Yang Feng et al. [15] have developed a SAW-based sensor for the measurement of the yarn tension and got the final result in terms of frequency shift with respect to the applied force. F. Wang et al. [16] have developed the humidity sensor based on the SAW delay line structure. They have linear variation in the frequency shift with respect to relative humidity. H. Weng et al. [17] have developed a SAW-based sensor for high-temperature measurement of the moving parts in the aero-engine system. L. Li et al. [18] have fabricated the SAW-based sensor for strain measurement in a high-temperature environment. They have collected the response for the proposed sensor in terms of frequency shift.

In the present work, we have proposed a SAW sensor-based wireless target type flow meter. Target type flowmeter is used in the industry frequently due to its many advantages. But the metering of the sensor is difficult as the sensor is placed inside the pipe section and metering with signal conditioning is outside the pipe section. Modifications on the target type flow meter have been proposed by different researchers to overcome this limitation. Modifications on the target type flow meter have been proposed by different researchers to overcome this limitation. Such modifications have been made by the researchers and are reported. Yong Zhao et al. [19] have fabricated a system by modifying the target type flowmeter with fiber brag grating, which is capable to measure the flow rate of liquid and gives the output in terms of wavelength shift. X. Qiao et al. [20] have modified the target type flow meter using the fiber Bragg grating and performed the experiments. The output of the flow of the liquid is proportional to the wavelength shift. Another modification approach on conventional target type flow meter has been made by C. Liu et al. [21] for the liquid flow measurement by modifying the conventional target type flow meter with the FBG. The output of the system provides a relation between the wavelength shift vs liquid flow rate. In the conventional target flowmeter, the force-balance type (target type) and strain gauge are used as primary and secondary sensors, respectively. The maintenance cost of the conventional-type target flowmeter is also very high compared to the SAW-type target flowmeter. As a result, the conventional target flowmeter becomes very much costly. The SAW-based target flow meter is useful for difficult measurements such as slurries, viscous or dirty fluids. Hence, the proposed flowmeter may be more rugged and accurate than the conventional target-type flowmeter. A conventional target flowmeter is used in the industry for sending the measured flow rate to a remote place through wires whereas the proposed one can do the same wirelessly which is very useful in hazardous areas and inflammable industries. In some previously modified target flow meter [19-21] uses FBG where the transmission of the sensor data can be done through optical channel which is very costly method.

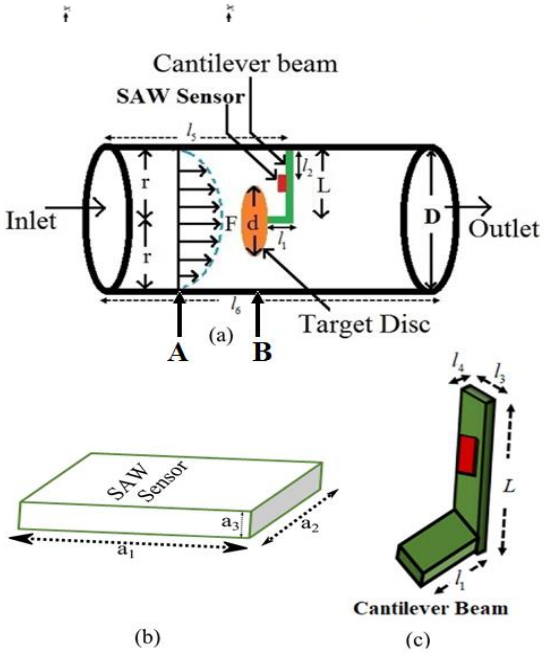


Fig. 1. Parameters used in modified target type flowmeter (a) Pipe configuration (b) SAW die configuration (c) cantilever beam with SAW die

II. METHOD OF APPROACH

A. Simulation of target rod flow meter with SAW Sensor

1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:18:54 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2021.3104863, IEEE Sensors Journal

The target type SAW-based flow sensor is shown in Fig.1. The arrangement is done in such a way that it is capable to transmit the measured information through the wireless domain with the help of SAW sensor. It consists of a pipe having a diameter of *D*. Inside the pipe, an *L* shaped vertical cantilever beam is fixed from the top of the pipe surface, which is perpendicular to the pipe axis. A disc having a diameter of *d* is mounted on the bottom end of the cantilever. SAW sensor is also mounted on the cantilever beam shown in Fig.1. When liquid flows through the inlet valve of the pipe and interact with the disc, the disc feels force *F* which is developed due to the flow of the liquid and contact with the surface area of the disc. The developed force is transferred to the cantilever beam with the help of disc, and the cantilever is displaced. A SAW sensor is fixed on the cantilever beam. The reference sensor also feels the force that generates the strain on the sensor layer.

Assuming the average flow rate of liquid in the pipe *V,* the flow velocity within the pipe at a random place can be written as

$$v(r) = 2V\left[1-\left(\frac{r}{0.5D}\right)^2\right] \tag{1}$$

where *D* is the diameter of the pipe and r is the radial distance from the centre of the pipe, for mechanical energy at places A and B is constant. So the kinetic energy generated due to flow of the liquid having velocity *v(r)* at point A is transferred into pressure increment at the point B is shown in Fig.1

$$\Delta p = \frac{\rho v(r)^2}{2} \tag{2}$$

where ρ is the density of the fluid.

So, the overall force *F* exerted on the target disc can be written as a product of the area of the target disc and change in pressure

$$F = \iint \Delta p.dr \tag{3}$$

where *dx* is the surface element of the target disc.

The force can be calculated on the disc having diameter *d* written as

$$F = \int_0^{d/2} 2\pi r \frac{\rho.v(r)^2}{2} dr \tag{4}$$

$$F = \frac{96\pi\rho V^2}{D}\left(48D^4d^2 - 12D^2d^4 + d^6\right) \tag{5}$$

Or $$F = K_1V^2 \tag{6}$$

where $$K_1 = \frac{96\pi\rho}{D}\left(48D^4d^2 - 12D^2d^4 + d^6\right) \tag{7}$$

For incompressible fluid flow, we can write velocity of liquid (V)

$$V = Q/A \tag{8}$$

where Q is liquid flow rate and A is the area of the pipe

Substituting the value of *V* in the equation (6).

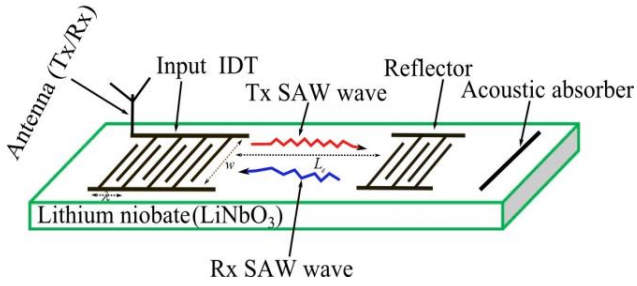


Fig. 3. 3D view of the SAW Delay line model with IDT and reflector

$$F = K_1\left(\frac{Q}{A}\right)^2 \tag{9}$$

$$Q = K_2\sqrt{F} \tag{10}$$

Equation (10) represents the relation between flow rate and the force developed in the target rod. The relation of strain with force is proportional and can be expressed as.

$$\varepsilon = \frac{6FL}{Ebh^2} \tag{11}$$

where *L* is the length of the cantilever, *h* is the thickness of cantilever, *E* is the young modulus of the material used in the cantilever, and *b* is the width of the fixed end of the cantilever. The Stress σ_x developed in the target rod in the *x*-direction for maximum rod displacement of t_1 from the reference position is.

$$\sigma_x = \frac{F\cdot x\cdot t_1}{J} \qquad (0<x<l) \tag{12}$$

The moment of inertia *J* of the rod can be written as

$$J = \frac{et_1^3}{12} \tag{13}$$

Strains developed along the beam due to flow in the x and y direction are ε_x and ε_y respectively and can be written as

$$\varepsilon_x = \frac{F.t_1.x}{E.J} \tag{14}$$

$$\varepsilon_y = \frac{\mu\cdot F\cdot t_1\cdot x}{E.J} \tag{15}$$

So the average strain distribution in the SAW die can be calculated in the direction of x and y are and $\overline{\varepsilon_x}$ $\overline{\varepsilon_y}$ respectively when the SAW die is fixed at the end of the cantilever.

$$\overline{\varepsilon_x} = \frac{\int_{d_1}^{d_2}\frac{F.t_1.x}{E.J}dx}{(d_1-d_2)} \tag{16}$$

$$\overline{\varepsilon_x} = \frac{F.t_1.x\cdot(d_1+d_2)}{2\cdot E.J} \tag{17}$$

$$\overline{\varepsilon_y} = \frac{\mu\cdot F.t_1.x\cdot(d_1+d_2)}{2\cdot E\cdot J} \tag{18}$$

where E is the Young's modulus and μ is the Poisson's ratio of the SAW die. As the strain changes, the resonant frequency

1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:18:54 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2021.3104863, IEEE Sensors Journal

Δf of the SAW sensor changes. Which is formulated as the expression,

$$\Delta f = \frac{(r_1 - 1)\bar{\varepsilon}_x + r_2\bar{\varepsilon}_y}{1 + \bar{\varepsilon}_x} f_0 \tag{19}$$

As the strain in the x direction is very small $(1 >> \varepsilon_x)$ therefore equation (19) can be simplified as

$$\Delta f = ((r_1 - 1)\bar{\varepsilon}_x + r_2\bar{\varepsilon}_y) \cdot f_0 \tag{20}$$

The deflection and stress level of the disc with the target rod at different force values is shown in Fig.2 (a) and (b); the respected colour contour is shown in Fig.2(c). The figure provides the information of different stress levels throughout the target length. By substituting the value of the equation (17) and (18) into the equation (20) yields:

$$\Delta f = \frac{6F}{Eet_1^2}[(r_1 - 1) - (r_2\mu)].f_0 \tag{21}$$

where r_1 and r_2 are the strain coefficients and f_0 is the reference resonant frequency. Putting the value of F from equation (9) to equation (21), we can write

$$\Delta f = \frac{6K_1Q^2}{Eet_1^2A^2}[(r_1 - 1) - (r_2\mu)].f_0 \tag{22}$$

$$K_2 = \frac{6}{Eet_1^2A^2}[(r_1 - 1) - (r_2\mu)].f_0 \tag{23}$$

Or,

$$\Delta f = K_1K_2Q^2 \tag{24}$$

Or,

$$\Delta f = K_3Q^2 \tag{25}$$

where $K_3 = K_1K_2$ (26)

$$\Delta t = K_4Q^2 \tag{27}$$

equation (25) shows the nonlinear relation between the flow rate and the frequency or time delay response. The output obtained from the computational post-processing unit is nonlinear with respect to the flow rate of the liquid as written in the equation (27). This non-linearity limits the desired output, and to obtain the desired output, artificial neural network algorithm is used. ANN-based functional block diagram is shown in Fig 5. The block diagram has the nonlinear input (Δt) and weights (W_{11}, W_{12} and W_{13}) obtained during training where TH and LN are Tan hyperbolic activation function and Linear activation function respectively. The performance is checked in ANN with different neurons (2, 3 and 5), and the algorithm used here is feed-forward back-propagation. The best linearized result obtained with three neurons. The final output after processing through the ANN block is written in equation (28).

$$Q = K_5\Delta t \tag{28}$$

where K_3, K_4 and K_5 are the constant terms.

B. SAW Sensor delay line Model

When a sinusoidal electrical signal is applied to the IDT port, it converts the electrical signal into the surface acoustic wave by developing the periodically distributed mechanical forces via the piezoelectric effect and generates the surface acoustic wave on the piezoelectric substrate shown in Fig.3 and the

Functional diagram for time delays computational Unit shown in Fig.4. The SAW wave is converted into an electrical wave and transmits the signal through an attached antenna. RF pulses are generated by SAW interrogation unit for the period 0 to 1.939×10^{-9} sec which passes through antenna and propagates through the medium (air), having a velocity of 3×10^8 *m/s*. When the antenna interacts with the received RF pulse, it converts the RF wave into the electrical signal which

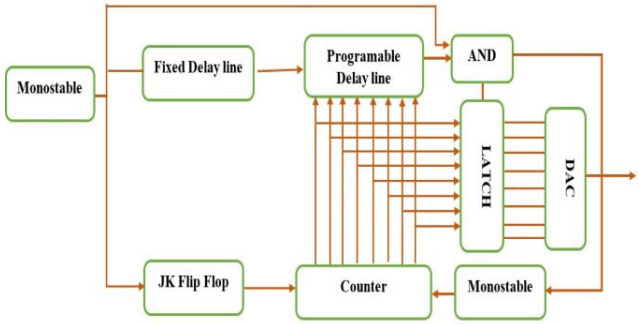


Fig. 4. Functional diagram for time delays computational Unit

further goes to the IDT. Now the IDT starts to convert the electrical signal into SAW (mechanical wave) and vice versa. The mechanical wave which is generated by the IDT on the Piezo electric substrate is known as Rayleigh wave.

1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:18:54 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2021.3104863, IEEE Sensors Journal

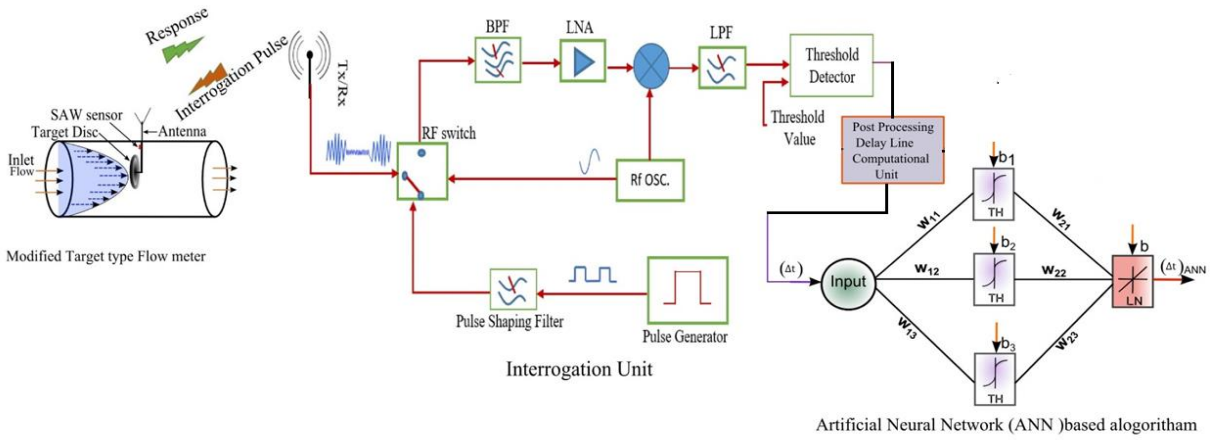


Fig 5. Functional diagram of the modified target type flow meter with SAW interrogation unit And Artificial Neural Network block

When the Rayleigh wave passes through the reflector, it provides some delay in the time domain depending on the change in applied pressure. This delay also depends on the physical properties of the piezoelectric substrate like dimension, density etc. The velocity of SAW also depends on temperature, i.e., the time impulses response of the reflected wave is also temperature-dependent. The temperature can be calculated with the help of the temperature coefficient T_k . The software-based temperature compensation program is already loaded in the interrogation unit. Now delayed mechanical wave again reaches to interdigitated transducer (IDT) and is transformed into an electrical signal. This electrical signal is then transmitted through SAW antenna and again reaches to SAW interrogation unit. The interrogation unit contains the low noise amplifier (LNA), demodulator and decoding components, which further process the signal.

When the liquid flows through the pipe and comes in contact with the target disc, a force is produced; due to this force, strain is produced. This produced strain is observed by SAW sensor. Because of the variation of the strain of the target rod, The SAW sensor also experiences stresses and strain that vary. This effect changes the density of the material, which affects the Rayleigh wave velocity. So, it generates a time delay between the reference signal and the received signal of 8ns. After detecting the received signal, the time delay response is extracted by the delay response analyzer circuit. The received signal passes to the monostable circuit, which triggers the J-K flip flop. The counter starts counting till the next pulse reaches the monostable vibrator. Digital to analog converter (DAC) which is connected with the LATCH used for simultaneous conversion from digital to analog output.

III. DESIGN

A wireless flow transducer is designed by modifying a conventional target type flowmeter with the help of a SAW sensor. In this, a PVC pipe is chosen to have a diameter D of 76 mm. The material used for developing the cantilever is Maraging Steel, whose young modulus E is $1.89 \times 10^5 N/m^2$. The cantilever beam having dimensions 38×2×2 mm is used for supporting the disc. For maximum interaction with the flow of the liquid, a disc is fixed in the middle of the pipe with the help of a cantilever, as shown in Fig.1 (a) in the schematic diagram. The radius of the disc is 15mm, and the thickness is 1mm. To observe the strain of the target, rod a SAW delay line model is chosen here. SAW sensor is mounted on the target rod near the pipe, as shown in Fig.1(c). A cantilever system is fixed with the pipe. Parameters with the 2D view of the SAW sensor delay line model is shown in Fig.10. To prevent the SAW sensor from direct contact with the liquid passivation and protection layer is used. SAW sensor wafer is a $128^{\circ} Y\ X$ cut (lithium niobate) crystal having Rayleigh wave velocity V_r of 3996 m/s. The dimension of SAW is chosen 12×1.5×0.5 mm. SAW die contains IDT for sensing and generating the SAW wave. The Material used for IDT fabrication is aluminium, having a thickness of 150 A°. IDT, along with the SAW die, has an operational frequency f_0 of 433 MHz. The adhesive (M-Bond 200 and M-Bond 200 catalyst) used for fixing the SAW sensor on the cantilever beam. The adhesive can withstand the temperature range [-30° to +65°C]. M-Bond 200 can be used for high elongation tests in excess of 60,000 macrostrain. To operate at this frequency inductor coil is developed on the printed circuit board PCB, is feed by a microstrip feeding technique.

TABLE I
CONFIGURATION OF MODIFIED TARGET TYPE FLOW MEASUREMENT SIMULATION PARAMETES DESIGNED ON COMSOL

Parameters	Description	Value
D	Diameter of the pipe	76 mm
l_6	Length of the pipe	110 mm
d	Diameter of the target disc	30 mm
l_2	Sensor position	5mm
l_1	Bend length	10 mm
L , l_3 , l_4	Cantilever beam dimensions	38 mm, 8 mm, 1mm
l_5	Position of the cantilever in the pipe from the inlet section	55mm
a_1 , a_2 , a_3	SAW sensor dimensions	12 mm, 6 mm, 0.5 mm

1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:18:54 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2021.3104863, IEEE Sensors Journal

TABLE II
SAW SENSOR CONFIGURATION WITH AND SIMULATION PARAMETERS

Parameters	Equation	Value	Description
f_0	NA	433 MHz	Applied Frequency
v_r	NA	3996 m/s	Rayleigh wave velocity
λ	v_r / f_0	9.22×10^{-6} m	Target wavelength
T_1	$\lambda / 4$	2.307×10^{-6} m	Electrode width
w	$5*\lambda$	4.614×10^{-6} m	Electrode length
N_1	NA	20	No. of electrodes i/p
N_2	NA	5	No. of electrodes Reflector
Pitch	λ	9.22×10^{-6} m	Pitch of electrodes
t_0	NA	50×10^{-6} m	Substrate thickness
V_0	V_{pp}	20 volts	i/p voltage magnitude
L_d	NA	16×10^{-6} m	Delay line expansion

The base material used for developing the antenna was an FR4 sheet having dimension (95mm x 45mm), and the thinness is 1.5 mm, designed on the PCB. The whole antenna circuit is passivated to prevent the external environment. The overall functional diagram of the proposed target type flowmeter is shown in Fig.5. The dimensions of the modified flowmeter with different parameter values are enlisted in Table I. The values of the different parameters and descriptions are given in Table II.

IV. Experimental and Simulation Result

To generate the response of the system for different parameters, parametric analysis, simulation is performed on the COMSOL software. The relation is calculated between force, and target rod displacement is shown in Fig.6. From the graph, it is observed that with the increase of force, displacement of the rod also increases. Here the free end of the rod is represented as 0.0 cm, and the fixed end is represented as 3.6 cm in Fig.7, which shows the displacement of the target rod at different positions of the rod for different flow rates from 0 LPH to 1200 LPH. One end of the rod is fixed, and another end is free, so when the force acts on the free end of the target rod, it behaves like a cantilever beam, and it develops different strain levels throughout the target rod shown in Fig.8. The strain developed in the rod is a maximum of 1.2×10^{-4} at the fixed end and a minimum of 9.236×10^{-6} at the free end when the flow rate is 1200 LPH. In the second phase, SAW die is mounted on the target rod so it senses the same force exerted by the liquid on the target rod. SAW die is rigidly fixed with the target rod near the pipe surface and is also displaced by certain values, as shown in Fig.9. When no force is act on the target rod, there is no displacement in SAW die. With the increase of force, due to an increase in flow rate, displacement of SAW die increases. As force reaches the maximum level of 0.926 N for the flow rate of 1200 LPH the SAW die is displaced by 1.8×10^{-3} cm. At 1200 LPH flow, the deflection is a maximum of 1.81×10^{-3} cm at the bottom of SAW die and minimum 2.29 × 10−4 cm at the top of the SAW die, as shown in Fig.10. The change in strain in SAW die with respect to the force is shown in Fig. 11. The strain developed throughout the length of the SAW die at different flow rates is shown in Fig.12. The shape is in the form

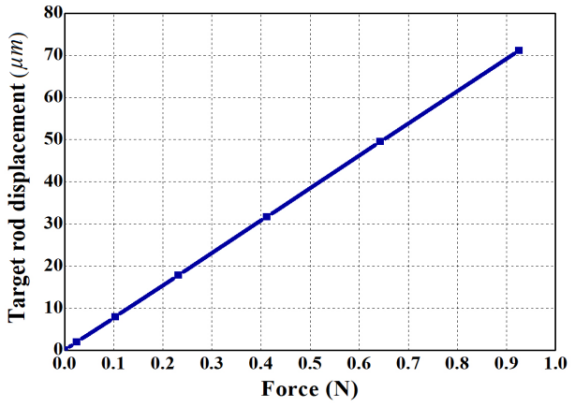


Fig. 6. Target rod displacement vs Force

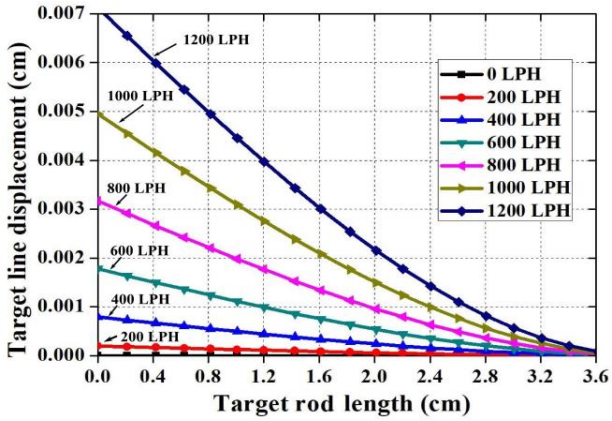


Fig. 7. Target rod displacement vs Target rod length on different flow rate

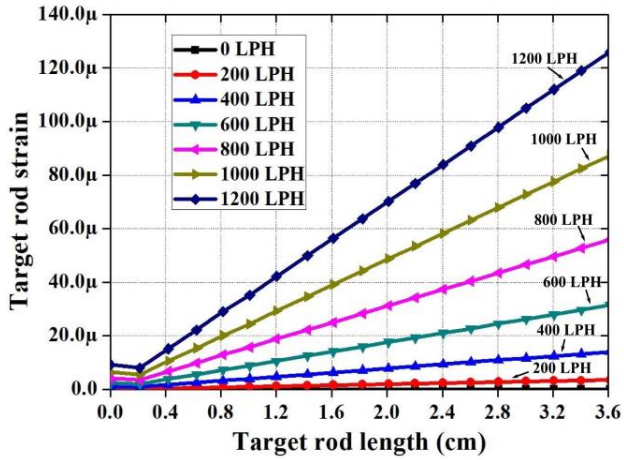


Fig. 8. Target rod strain vs target rod length on different flow rate.

of the inverted U, in which the characteristics are almost linear between the rate and target rod displacement. Displacement of target rod and strain on the SAW die with respect to the flow rate of liquid are shown in Fig.13 and Fig.14.The interrogation signal is given to the input IDT shown in Fig.15 (a). The voltage level of the interrogation signal is *V pp* 20 Volt, which initially starts for the period of 0 to 1.939×10^{-9} sec. and goes down to zero for the period 2.0323×10^{-9} sec. to 4.138×10^{-9} sec. As the flow rate increases, the strain also increases, as shown in Fig.20. Due to change in strain of the SAW sensor

1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:18:54 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2021.3104863, IEEE Sensors Journal

with the flow rate, the different delayed response is generated which is shown in Fig.15 (b) to (f). The response of the delayed signal with respect to the reference signal is shown in Fig.16. The signal comes in the output at 0 LPH is named as the reference signal, which is used for the measurement of the delayed response of the signal. Time response analysis is calculated for the output signal with respect to the flow rate is shown in Fig.17 from the graph; it is clear that with the increase of the flow rate, output signal response is delayed. Each delayed o/p response contains the information of different flow rates. The graph shows the comparison output for different neurons in ANN and without ANN. The most linearity was obtained by using ANN with three neurons.

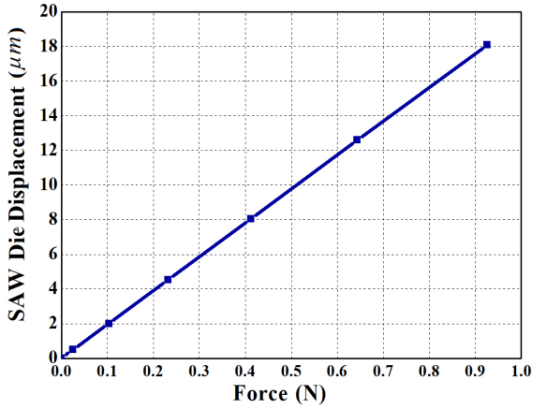


Fig. 9. SAW die displacement vs Force

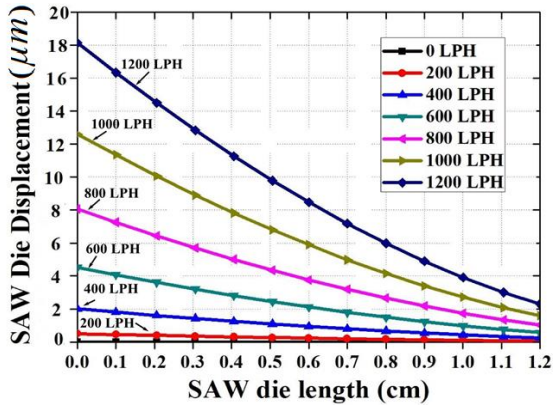


Fig. 10. SAW die displacement vs SAW die length on different flow rate

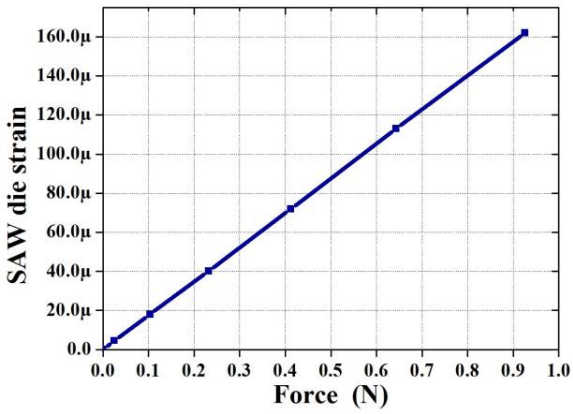


Fig. 11. SAW die with respect to the Force

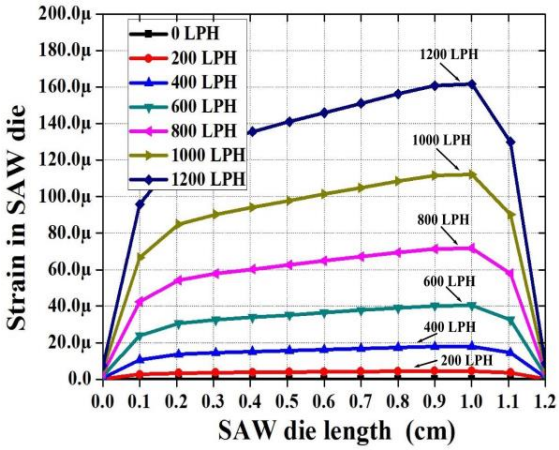


Fig. 12. Strain in SAW die with respect to the SAW die length on different flow rate.

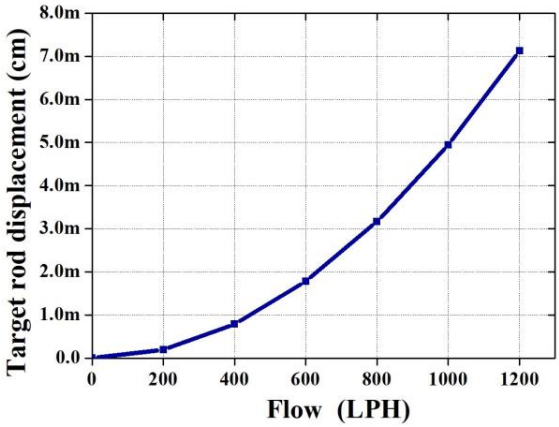


Fig. 13. Target rod displacement with respect to Flow rate

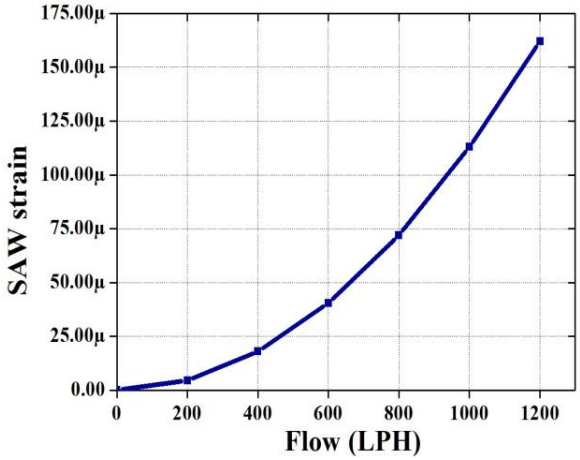


Fig. 14. Strain in SAW die with respect to Flow rate

1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:18:54 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2021.3104863, IEEE Sensors Journal

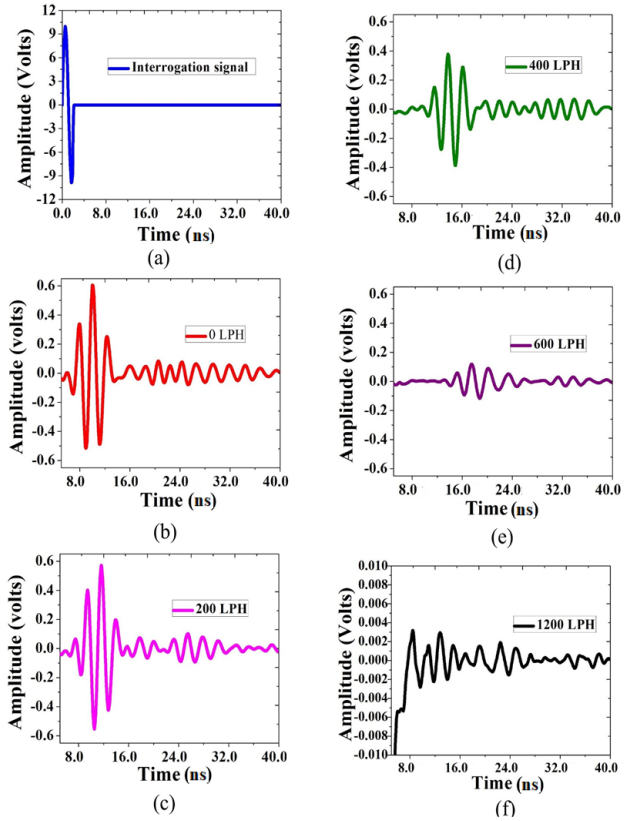


Fig. 15. (a) Interrogation signal. The received response(b) Amplitude vs time at flow rate 0 LPH, (c) amplitude vs time at flow rate 200 LPH, (d) Amplitude vs time at flow rate 400 LPH, (e) Amplitude vs time at flow rate 600 LPH, (f) Amplitude vs time at flow rate 1200 LPH

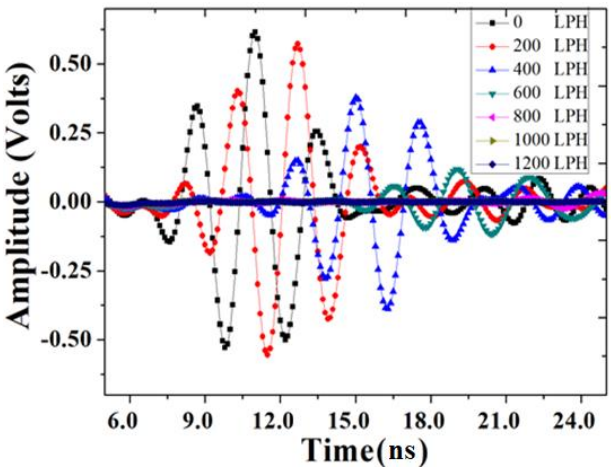


Fig. 16. Response of the signal vs time at different flow rate

V. Discussion

The target type flow measurement system is based on SAW technology. This technology offers the battery less operation. The system is very sensitive for small flow rate. The displacement of the target rod is linear with respect the force and it varies from 0 to 7 $\times 10^{-3}$ cm as shown in Fig.1.

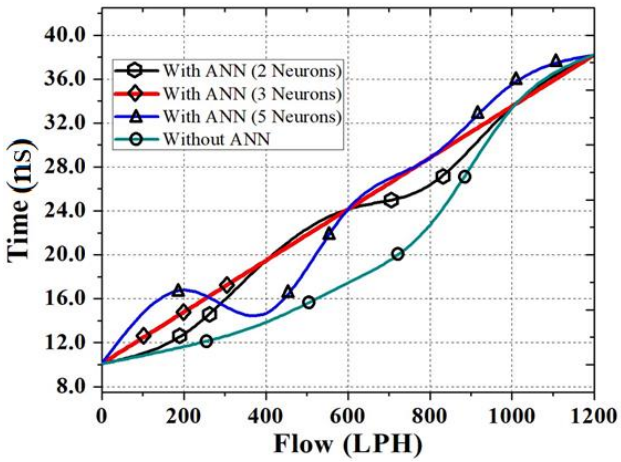


Fig. 17. Output response in time delay signal vs Flow rate using ANN with different Neurons

The displacement of the SAW die is also linearly varied with respect to the force, as shown in Fig.9. The displacement of the SAW die is maximum from its rest position by an amount of 1.8×10^{-3} cm as shown in Fig.10. Strain of the SAW die is also linearly varied with respect to the force, as shown in Fig.11. The maximum displacement of SAW die (1.8×10^{-3} cm) comes at the flow rate of 1200 LPH. Therefore, this sensor can be used to measure the fluid flow rate from 0 to 1200 LPH. Fig.15 deals with the response of the interrogation pulse for the different flow rates. The response is obtained in terms of a time-delayed signal having an amplitude value at the Y-axis. The system's output response is calculated in terms of a time-delayed signal which varies from 10×10^{-9} sec to 38.24×10^{-9} sec, as shown in Fig.17. It is also possible to enhance the sensor sensitivity and flow measurement range by shifting the SAW sensor's position. The output response of the sensor with flow rate is nonlinear which limits the response of the system, but system output can be linearize using different linearizing techniques. To overcome this problem, the proposed system uses the Multilayer Backpropagation ANN for linearization. Basically, for this system, ANN has been used for the linearization of the output response. The proposed method is static in nature. The flow of fluid flowing through the measuring pipeline section is static fluid. Hence the proposed neural network is compatible with this system. The nonlinearity linked with the proposed system often produces problems in readout or control applications. The environmental effects, manufacturing tolerances, and drift in performance due to ageing cause a measurement error; sometimes it changes the characteristics and shortens the life span of the proposed system. Hence, it causes the need for frequent calibration. And for the calibration, the ANN is one of the best techniques which reduce cost and time. It is a self-calibrated process that increases the accuracy and increasing the life span of the system. ANN offers a reliable tool that can solve, model, and predict a complex problem. ANN can solve a complex non-linear problem without knowing

1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:18:54 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2021.3104863, IEEE Sensors Journal

Table III
COMPARATIVE STUDY OF VARIOUS TYPE LIQUID FLOW MEASUREMENT TECHNIQUES

S. No	Technique	Range	Sensitivity	Linearity	Feature
1.	Y. Zhao et. al. [19]	0-3600 LPH	0.403 pm/LPH	±28.6% of FSO	FBG based target flow meter
2	X. Qiao et al. [20]	720-3600 LPH	0.868 pm/LPH	±13.5% of FSO	FBG based target type flow meter
3.	C. Liu et. al. [21]	0-7920 LPH	0.442 pm*/LPH pm -Pico meter	±15.7% of FSO	Target type flow meter on FBG sensing
4.	Proposed method	0-1200 LPH	0.0225ns/LPH	±3.703% of FSO	SAW sensor-based target type flow meter

the mathematical relationship between input and output. In particular, ANN is trained to function in definite environment conditions and can be easily retrained to deal with minor changes in the operating environmental conditions. The ANN- modeling approach incorporates intelligence into a proposed system to make it free from measurement error, environmental drifts, and ageing, etc. The ANN-based technique has certain advantages over classical regression and statistical methods. Traditional methods such as analogue, digital and computer look-up ROM tables; these methods are time-consuming tasks because calibration is done by manual and repetitive identification. These methods do extensive mathematical modelling and require a complete mathematical relationship between the independent and dependent variables, which is difficult for the complex non-linear system. The Neural network models require less statistical analysis. The Neural network models can indirectly detect complex non-linear relationships between independent and dependent variables. Neural network can easily implement in hardware form, and it has the potential to be inherently fault-tolerant. It can work well in case of incomplete information regarding the system. Thus, neural networks may represent a smart alternative to a statistical modelling technique. After linearization, the time delay response is proportional with respect to flow rate.

VI. Conclusion

In this paper, flow measurement and transmission systems have been designed based on SAW sensors. This transmission system is wireless and is capable of transmitting the signal to remote locations. In a conventional target type flowmeter, the force-balance type (target type) is used as a primary transducer, and a strain gauge is used as a secondary sensor. A comparative study of various types of liquid flow measurement techniques is shown in Table III. The proposed system is very sensitive with respective flow rate in comparison to the ref. [19, 20, and 21]. This system has special features; it does not require on board power supply, which increases the life cycle of the proposed system. This system is externally Rf powered. The advantages of SAW-based target flowmeter are that the information is send wirelessly, and very much suitable for hazardous and inflammable industries. It has no moving part, so the mechanical error is negligible. The maintenance cost of the SAW-based target flowmeter is also very low compared to the conventional type target flowmeter. Hence, it is more rugged and accurate than the conventional target type flowmeter.

References

[1] D. M. Considine, Process Instrument and Control Hand Book, 2nd ed. New York, NY, USA: McGraw-Hill, 1999.

[2] E. O. Doeblin, “Measurement System Application and Design,” 5th ed., York, NY, USA: McGraw-Hill, 2000.

[3] S. Sinha, D. Banerjee, N. Mandal, R. Sarkar and S. Chandra Bera, “Design and Implementation of Real-Time Flow Measurement System Using Hall Probe Sensor and PC-Based SCADA,” *IEEE Sensors Journal*, vol. 15, no. 10, pp. 5592 – 5600, Oct.2015.

[4] Clemens C.W. Ruppel, and Tor A. Fjeldly, “Wireless SAW identification and sensor systems,” in Advances in Surface Acoustic Wave Technology, Systems And Application 1st ed., Singapore, World Scientific, 2001, ch. 4th, pp. 270 – 280.

[5] A. Mason, Subhas C. Mukhopadhyay, K. P. Jayasundera and N. Bhattacharyya, “High frequency Surface Acoustic Wave (SAW) Device for Toxic Vapor Detection,” in Sensing Technology: Current Status and Future Trends II 1st ed., Switzerland, Springer, 2014, ch.11, pp. 221 – 224.

[6] S. G Joshi, “Surface-Acoustic-Wave (SAW) Flow Sensor,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 38, no.2, pp. 148 – 154, March 1991.

[7] W. E. Bulst, G. Fischerauer and L. Reindl, "State of the art in wireless sensing with surface acoustic waves," in *IEEE Transactions on Industrial Electronics*, vol. 48, no. 2, pp. 265-271, April 2001.

[8] T. Moazzeni, J. Ma, Y. Jiang and N. Li, “Flow Rate Measurement in a High-Temperature, Radioactive, and Corrosive Environment,” IEEE Transactions on Instrumentation and Measurement, vol. 60, no. 6, pp. 2062 – 2069, June. 2011

[9] R. Stoney, B. Donohoe, D. Geraghty, G.E. O’Donnell, “The development of surface acoustic wave sensors (SAWs) for process monitoring”, 5th CIRP Conference on High Performance Cutting 2012, p 569 – 574

[10] Y. Wang, Z.Li, L. Qin, M. K. Chyu and Qing-Ming Wang, “Theoretical and Experimental Studies of a Surface Acoustic Wave Flow Sensor,” IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 59, no. 3, pp. 481 – 490, March 2012.

[11] V. Kalinin, “Modelling of Hysteresis and creep in SAW strain sensors”, Proc. of 2014 Joint Conference of the IEEE International Frequency Control Symposium & European Frequency and Time Forum, May 19-122, 2014 Taipei, Taiwan, pp. 481-484

[12] V. Kalinin, A. Leigh and A. Stopps, “Stability and durability of resonant SAW strain sensors” Proc. of 2016 European Frequency and Time Forum, Apr. 4-6, 2016, York, UK.

[13] V. Kalinin, “Wireless passive resonant SAW sensor for monitoring temperature, strain, torque and pressure”, Proc. of VII ECCOMAS Thematic Conference on Smart Structures and Materials SMART 2017, 6-8 June 2017, Madrid, Spain, p. 1453-1464

[14] L. M. Rodriguez, D. R. Gallagher, M. W. Gallagher, B. H. Fisher, J. R. Humphries, and D. C. Malocha, “Wireless SAW sensor temperature extraction precision,” IEEE Sens. J., vol. 14, no. 11, pp. 3830–3837, 2014.

[15] Y. Feng, Z. Lu, W. Lu, C. Zhu, Q. Liu and H. Zhang, "Study of the Doubly Clamped Beam Yarn Tension Sensor Based on the Surface

1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:18:54 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2021.3104863, IEEE Sensors Journal

Acoustic Wave," in IEEE Transactions on Industrial Electronics, vol. 66, no. 4, pp. 3256-3264, April 2019.
[16] F. Wang et al., "A Miniaturized Integrated SAW Sensing System for Relative Humidity Based on Graphene Oxide Film," in IEEE Sensors Journal, vol. 20, no. 17, pp. 9733-9739, 1 Sept.1, 2020.
[17] H. Weng, F. L. Duan, Z. Xie, S. Liu, Z. Ji and Y. Zhang, "LiNbO3-Based SAW Sensors Capable to Measure up to 1100°C High Temperature," in IEEE Sensors Journal, vol. 20, no. 21, pp. 12679-12683, 1 Nov.1, 2020.
[18] L. Li, B. Peng, J. Zhu, Z. He, Y. Yang and W. Zhang, "Strain Measurements With Langasite SAW Resonators at High Temperature," in IEEE Sensors Journal, vol. 21, no. 4, pp. 4688-4695, 15 Feb.15, 2021.
[19] Y. Zhao, K. Chen and J. Yang, "Novel target type flowmeter based on a differential fiber Bragg grating sensor," Measurement –Elsevier, vol.38 pp. 230 – 235, 2005.
[20] X. Qiao, Q. Zhang, H. Fu, and D. Yu, "Design of the target type flowmeter based on fiber Bragg grating and experiment," Chinese Opt. Lett., vol. 6, no. 11, pp. 815–817, 2008.
[21] C. Liu, Z. Zhang, H. Li, Z. He, X. Zhao, and R. Wu, "Research on OnePiece Structure Target Flow Sensing Technology Based on Fiber Bragg Grating," *Photonic sensors*, vol. 6, no. 4, pp. 303–311, 2016.



Praveen Maurya was born in Lucknow, India, in 1989. He received the B.Tech. degree in Electronics and Communication Engineering from Uttar Pradesh Technical University, Lucknow, India, in 2011, and the MTech. and Ph.D. (Eng.) degree in Electronics and Communication Engineering from IIT (ISM), Dhanbad, India, in 2014 and 2021, respectively. He has authored Five papers in international journals and eight international conferences. His current research interests include transducer development, Instrumentation, SAW sensor and MEMS.



Sayyed Faizan Ali was born in Kanpur, India, in 1989. He received B.Tech. degree in Electronics and Instrumentation Engineering from Uttar Pradesh Technical University, Lucknow, India, in 2011, and the M.Tech. degree in Control and Instrumentation Engineering from Delhi Technological University (formerly DCE), New Delhi, India, in 2015. He is currently pursuing the Ph.D. degree with the Department of Electronics Engineering, IIT (ISM) Dhanbad, Dhanbad, India. His current research interests include sensor design, controller design, and MEMS. He is also a member of the Optical Society of America and the Society of Photo-Optical Instrumentation Engineers.



Nirupama Mandal (M'15–SM'19) received the B.Sc. (Hons.) degree in physics, the B.Tech. degree in Instrumentation Engineering, the M.Tech. degree in Instrumentation and Control Engineering, and the Ph.D. (Tech.) degree in Instrumentation Engineering from the University of Calcutta, Kolkata, India, in 2001, 2004, 2006, and 2012, respectively. She has joined the Department of Electronics Engineering, IIT (ISM) Dhanbad, India, as an Assistant Professor in 2013 and currently working as an Associate professor. She is involved in carrying out research projects on transducer development, process-plant instrumentation, and process modeling. She has published over 50 papers in international journals and 35 articles in international and national conference proceedings. Her current research interests include transducer design, PC-based instrumentation, controller design, MEMS, SAW sensor, wireless transmitter, optical transmitter, and process modeling.

1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:18:54 UTC from IEEE Xplore. Restrictions apply.

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

Design and Performance Analysis for Underground Trains using Optisystem

Anurag Chauhan
Assistant Professor
Electronics and Communication
Delhi Technological University
anuragchauhan@dtu.ac.in

Rishabh Verma
Student
Delhi Technological University
rishabh.verma1603@gmail.com

Ashish Joshi
Student
Delhi Technological University
ajk98a60@gmail.com

Rishabh Sharma
Student
Delhi Technological University
rishabhsharma9971@gmail.com

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498461

***Abstract*—In this study, Wireless Optical Communication LOS model for underground trains to station based on FSO is designed and simulated to achieve optimum performance and reduced bit error ratio. The FSO medium is highly dependent on Air Quality and Weather conditions bit error ratio vary depending on the conditions. The aim of this design is to achieve a lower Bit error ratio and to increase the range for the communication. The simulation is designed to improve the quality of communication through FSO channel by reducing the bit error ratio through filtering techniques. The results of this simulation will help future research and implementations of the FSO based wireless optical communication for underground trains. This simulation is based on V2I architecture where the communication is based on FSO technology.**

***Index Terms*—FSO, Wireless Communication, Optical Communication, Bit Error Ratio, V2I**

I. Introduction

With the substantial growth of technology the communication technologies have significantly improved providing a faster and reliable end-to-end communication to consumers. The demand of more reliable services have greatly soared with the rapidly changing technology.In the midst of the uprising demand of bandwidth in metropolitan cities catering gigantic number of users created dearth of bandwidth spectrum for communication providers.To cater the demands of customers telecom providers drive the development of communication technologies while pursuing variety of methods to provide faster and reliable communications [1].

For communication, a range of wired and wireless communication systems are being used. Different parameters such as range, bandwidth, and transmission losses are used to classify these technologies. Each category of these technologies are used for a multitude of applications.The cost, range and bandwidth parameters have a direct influence on the applications of these communication technologies. Fiber optic cables are the first and often the most apparent option for wired communication technology, but the costs for laying fibre and resulting delays often make them economically prohibitive. Radio frequency (RF) technology option is a more mature wireless communication technology than Free Space Optics (FSO) for long distance communications but its restricted bandwidth capabilities are the drawbacks of RF technologies [1].

For digital communications over long and medium distances, optical fibre transmission has been the prominent technology over the past decade. Whereas for short distances, there are certain technical demerits of optical communication. As a result, optical communication did not perforate into fields such as board-to-board and chip-to-chip communication. A certain tie is shared between FSO and fibre-optic cable. Essentially, the principle of FSO and fibre optic communication are the same. The underlying principle related to optical communications through fibre-optic cables is the use of lasers. The main distinction between them is the medium [1].

The FSO technology is currently highly pursued for its numerous uses, such as terrestrial communication linkages, inter-satellite connections, ground-to-ground communication, air-to-ground communication, ground-to-air communications, for example Unmanned aircraft units, fibre backup, emergency corridors, transportation networks and last mile connectivity [2,4]. Even though FSO communication encompasses all of the advantages listed, atmospheric channel inconsistencies significantly hinder the performance of the communication. Because the FSO channel operates in the troposphere, where great deal of the atmospheric phenomena happens, leading to unpredictable poor weather conditions such as rain, haze, fog and drizzle, the availability and durability of FSO are adversely affected[2].

The research and development institutions are now focusing on this vehicle communication technology progressively. Driver support systems also known as Driver assistance Systems gather data from different vehicle-integrated sensors to warn drivers of potential dangers and enhance driving efficiency. However, only in a small area can the sensors used in the vehicle record data. The connection between Vehicles to Vehicles (V2V) and Vehicles to Infrastructure (V2I) will increase driver's knowledge of potential dangers and threads. FSO is the infrastructure that supports future automotive network's high bandwidth demands[4,5].

II. Objective

In the previous researches, FSO system was designed and simulated using Optisystem for underground trains in both idle and moving states[3]. To have a secure transmission, narrow laser beam is used to cover the tracks of the train.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:57:59 UTC from IEEE Xplore. Restrictions apply.

The simulation was performed and analysed for a distance of 1400m where the results of bit error ratio, Max Q-Factor and Range were taken into consideration. The maximum track coverage with bit error rate equals zero was 800m for bitrate of 1.25Gbps. This research simulate FSO communication model for underground trains to improve bit error ratio and Max Q Factor of the received signal as well as improve the range coverage of the system by using filtering techniques[3].

III. Proposed Method

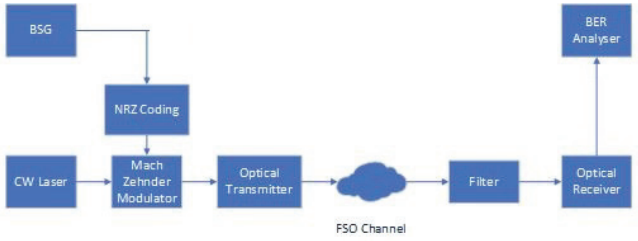


Fig. 1. Communication Block Diagram

This work is performed for bit rates of 1.25 Gbps and have analysed the relationship of Max Q-Factor and BER with Range. The simulation design parameters used in the work are shown in Table 1. The FSO transmitters commercially available are broadly classified in two categories based on the wavelengths of the transmitted laser. One operates on 800 nm wavelength and the other one operates on 1550 nm. The Eye safety standards and regulations permits laser transmitter that operates on wavelength of 1550 nm. This transmitter operating on 1550 nm can transmit 50 times the power than can be transmitted on 800 nm wavelength.Hence, for simulation the laser source of 1550 nm wavelength was used for transmitting the information from source to destination[3].

IV. Experimental data and results

For this V2I simulation model continuous wave laser is used for transmission and the information or data bit sequence is generated using a pseudo random bit sequence generator (PRBSG) where the bit rate of the sequence generator is configured for 1.25 Gbps to simulate and obtain the BER and Q-Factor for performance evaluation. CW laser was configured for frequency of 193.1 THz with optical power of 25 dBm. The signal generated by the PRBSG was converted to an electrical signal using a non return to zero (NRZ) pulse generator. The NRZ signal was then modulated on the continuous wave laser through mach zehnder modulator. The output of mach Zehnder modulator is an optical signal which will be transmitted through the FSO channel. The transmitter aperture diameter used for transmission is 5cm whereas the receiver aperture diameter is 20 cm and having beam divergence of 2 mrad. The transmitter loss was configured as 2 dB for the simulation and the attenuation of the FSO channel was considered as 25 dB/km.

The configured range of the FSO channel was from 650m to 1190m for which the values of BER and Q-Factor were observed. It was configured in sweep mode for simulation to determine the relationship of BER vs Range and Max Q-Factor vs Range for the above-mentioned bit rate. The Range parameter was set for 19 iterations which will cover a total distance of 650m to 1190m (i.e., 30 m for each iteration). The responsivity of the receiver used for simulation is 1 Ampere per Watt.The simulation was performed on the basis of the design parameters shown in table 1.

TABLE I
Simulation Parameters [3]

Parameter	Value
CW Laser Transmitter	
Frequency	193.1 THz
Optical Wavelength	1550 nm
Optical Power	25 dBm
FSO Channel	
Range	0-1400 m
Attenuation	25 dB/km
Beam Divergence	2 mrad
Transmitter aperture diameter	5 cm
Receiver aperture diameter	20 cm
Transmitter loss	2 dB
Receiver loss	0 dB
Optical Receiver	
Responsivity	1 A/W
Pseudo Random Bit Sequence Generator	
Bit Rate	1.25 – 4.25 Gbps

A. Simulation Results

TABLE II
BER vs Range

	Bit Error Ratio	
Range(m)	Without Filter	With Filter
650	0	0
680	4.87E-274	2.67E-283
710	8.88E-171	1.62E-192
740	1.81E-106	1.04E-125
770	1.71E-66	1.21E-80
800	1.01E-41	1.68E-51
830	1.99E-26	4.22E-33
860	4.98E-17	1.65E-21
890	3.07E-11	3.36E-14
920	1.16E-07	1.37E-09
950	1.99E-05	1.15E-06
980	0.000513	8.38E-05
1010	0.004102	0.001301
1040	0.01891	0.007648
1070	1	1
1100	1	1
1130	1	1
1160	1	1
1190	1	1

The results of the simulation were calculated through optisystem 7 software. The outputs of the simulation design was measured through BER analysers. We measured the output for two communication designs one without filter and one with a low pass gaussian filter having cut off frequency of

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:57:59 UTC from IEEE Xplore. Restrictions apply.

7.5 GHz. The output results of the simulation are shown in table 2 which represents the outputs of bit error ratio and table 3 which represents the outputs of Max Q Factor for both with filter and without filter simulation designs.

TABLE III
Q FACTOR VS RANGE

	Max Q Factor	
Range(m)	Without Filter	With Filter
650	44.865273	42.431862
680	35.346814	35.945499
710	27.827135	29.570235
740	21.878838	23.815172
770	17.177805	18.975675
800	13.471488	15.044164
830	10.560573	11.91813
860	8.289752	9.440101
890	6.520087	7.47716
920	5.144315	5.925727
950	4.074226	4.698486
980	3.239328	3.73016
1010	2.586711	2.967738
1040	2.007544	2.368677
1070	0	0
1100	0	0
1130	0	0
1160	0	0
1190	0	0

From the results of the simulations the relationship between the range and BER and range and Q factor have been obtained which are shown in figure 2 and figure 3 respectively.

B. BER Vs Range Relationship

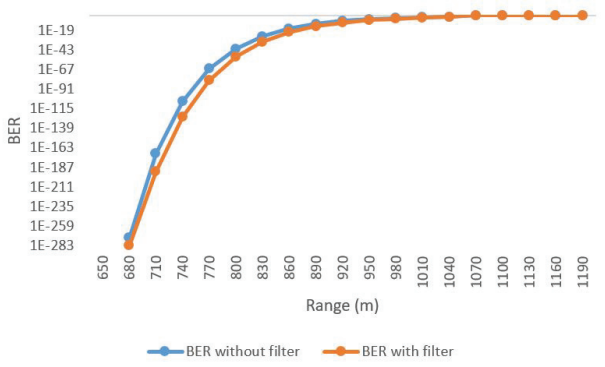


Fig. 2. BER Vs Range

From the figure 2 and table 2 it can be observed that the bit error ratio for the simulation design consisting of filter is less than the simulation design without filter. As per the standards of communication when a BER of less than 1E-9 is achieved then the received signal is considered to be a good signal. But when the BER is greater than 1E-9 then the communication is categorised as poor communication and the received signal consist of errors. Following these standards the signal received with a BER nearing 1E-9 (i.e. 1.37E-09) is received at a distance of 920m for simulation design consisting low pass gaussian filter whereas at 890m (i.e. 3.07E-11) for simulation design without filter. The signal received after this distance have BER greater than 1E-9 and hence those signals are not usable to successful communication.

C. Q Factor Vs Range Relationship

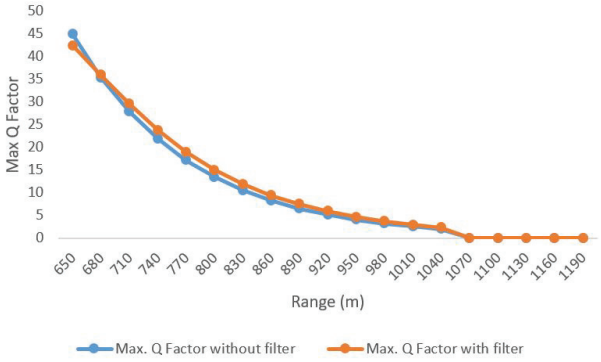


Fig. 3. Q Factor Vs Range

From the figure 3 and table 3 it can be observed that the Q Factor for the simulation design consisting of filter is more than the simulation design without filter. As per the standards of communication when a Q Factor of more than or equals 6 is achieved then the received signal is considered to be a good signal having sufficient signal quality. But when the Q-Factor is less than 6 than the communication is categorised as poor communication and the quality of received signal is insufficient. Following these standards the signal received with a Q Factor of nearing 6 (i.e. 5.925727) is received at a distance of 920m for simulation design consisting low pass gaussian filter whereas at 890m (i.e. 6.520087) for simulation design without filter. The signal received after this distance have Q Factor less than 6 hence those signals are not usable to successful communication.

From the results of the simulations and the observations it can be implied that the results of simulation design using the low pass gaussian filter gave more promising results than the simulation design without filter.

V. CONCLUSION

On the basis of the results of simulation and relationship graphs between BER versus Range and Q factor versus Range it can be concluded that the performance achieved by the communication model consisting of filter is significantly better than the communication model without filter. It can also be stated that the FSO based communication between underground trains can be improved by using a low pass gaussian filter at the receiver section. The performance of the communication model consisting of filter can be further improved by using encoding and error correction techniques.

VI. REFERENCES

1) Sadiku, M and Musa, S and Nelatury, Sudarshan,"Free Space Optical Communications:

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:57:59 UTC from IEEE Xplore. Restrictions apply.

An Overview", *European Scientific Journal*,2016,doi: 10.19044/esj.2016.v12n9p55.
2) Gurjit Kaur, Disha Srivastava, Prabhjot Singh, Yaman Parasher, "Development of a novel hybrid PDM/OFDM technique for FSO system and its performance analysis", in *Optics & Laser Technology*, vol. 109, 2019, pp. 256-262, ISSN 0030-3992, doi: https://doi.org/10.1016/j.optlastec.2018.08.008.
3) M. Sivaranjani and Dr. J. Vidhya, "System Performance of Free Space Optics in Underground Moving Train Using Optisystem 14". in *Asian Journal of Applied Science and Technology*,2017,pp. 232-235.
4) Aminian, Mohammad & Dong, Yao, "Routing in Terrestrial Free Space Optical Ad-Hoc networks", 2014, doi: 10.13140/2.1.2675.9369.
5) e. a. Y. Takayama, "Data transmission performed from a moving vehicle," in *SPIE Newsroom Astronomy*, 2013.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:57:59 UTC from IEEE Xplore. Restrictions apply.

Journal of the Iranian Chemical Society
https://doi.org/10.1007/s13738-021-02370-1

ORIGINAL PAPER

Design, synthesis and self-assembly of amide-linked dendron-based non-ionic amphiphiles

Ashwani Kumar[1] · Ram Singh[2] · Yogesh Kumar Tyagi[1]

Received: 6 May 2021 / Accepted: 3 August 2021
© Iranian Chemical Society 2021

Abstract
In this manuscript, we discussed the synthesis of non-ionic (C_{18}) and (C_{18}-cis) amphiphiles from [PG]-dendron-based biocompatible starting materials. The physicochemical studies of amphiphiles show some significant results. The presence of an amide as a linker group has affected the hydrophilicity of the amphiphiles to respond to a low critical micelle concentration. The graph of surface tension versus concentration shows that the CMC of non-ionic amphiphiles falls in the range of 18.10 µmol L^{-1}–00.61 µmol L^{-1}. The dynamic light scattering study of the synthesized compounds confirmed that the G.n-amide amphiphiles with the unsaturation 6, 8 and 10 have an average aggregated particle size range of 93.0 nm–13.0 nm, while the average aggregated particle size of saturated amphiphiles 13, 15 and 17 showed a range from 28.0 nm to 7.5 nm. The morphological study of self-aggregated particles of amphiphiles using transmission electron microscopy studies in an aqueous medium showed the worm-like shape for G.1 amphiphiles and circular shape for G.2 and G.3 amphiphiles.

Keywords Amphiphilic molecules · Physicochemical properties · Dendron · Micelle · Self-assembly

Introduction

The dendritic amphiphiles have shown their application in various fields [1–4]. The dendron-based amphiphiles are used in synthesis because of their simplicity [5–7], self-organizing [8–10] and unrivaled physicochemical properties [11–13]. Novel biodegradable dendritic amphiphiles that use a viable methodology and biocompatible starting material [3, 14] are an emerging research topic [15–20]. The development of dendritic non-ionic amphiphiles has received a lot of attention in nanoscience due to its extraordinary basic properties such as a larger number of branches and numerous head groups at the edge [21, 22]. They also show some superior aggregation properties in the bulk state, which has been the inspiration for further studies in aqueous solution. Since the subsequent structures have high stability [18], low critical micelle concentration [8, 23] and low Krafft temperature [24, 25], the dendritic amphiphiles are much advantageous than other nanostructures because of their high aggregation properties. Dendritic self-assembling structures also have a high proportion of useful groups on the periphery. They can be modified by various functionalities such as ligands, targeting molecules and other desired groups so that they are multivalent. They can imitate interacting Web sites, which is deeply desirable [26]. Dendritic amphiphiles with micellar scaffolds are a new method of drug delivery. The improvement in the bioavailability of targeted drugs can develop beneficial effects in a controlled manner over a longer period of time [27, 28]. Some dendritic self-assembling amphiphiles have limitations such as instability at certain pH values [29, 30]. There have been reports on the types of arrangement of amphiphiles, their arrangement parameters, factors influencing their behavior, their use in various arrangements and their therapeutic applications [31–33].

Due to the biocompatibility of amide groups, it has received considerable attention in the form of pharmaceutical compounds, biologically active components, polymers, surfactants and natural substances. The amide bond present in synthesized amphiphiles is biodegradable [28]. This bond is also known for its higher pH stability than the previously synthesized ester amphiphiles [6]. Unlike previously synthesized ester amphiphiles, amide-linked

✉ Yogesh Kumar Tyagi
drytyagi@ipu.ac.in

[1] University School of Basic and Applied Sciences, Guru Gobind Singh Indraprastha University, Dwarka, New Delhi, India

[2] Department of Applied Chemistry, Delhi Technological University, Rohini, New Delhi, India

Published online: 16 August 2021

dendron-based non-ionic amphiphiles reduce the CMC value [13]. In contrast to a saturated alkyl chain, the unsaturated alkyl chain interacts with hydrophobic drug or dye molecules with electrons to give stability to the host and guest assembly.

Here, we have expanded our work from the ester bond to the amide bond to better understand the change in physicochemical properties. The synthesis, characterization and physicochemical properties of the non-ionic amphiphiles on amphiphiles based on amide-bonded dendrons were achieved. The data in Table 1 show that the amides, in contrast to esters, are very stable in an aqueous solution. A comparison of the physicochemical and pH sensitivity of amide in conjunction with an ester counterpart previously described showed a significant reduction in the various surface activity of the amide. This corresponds to a decreasing CMC, particle size and polydispersity index (Table 1). This observation indicates that amides have stronger intermolecular interactions and are more polar than the ester. Compared to the ester amide, they are very stable and have a longer half-life. The relatively slow hydrolysis of amide over esters in an acidic environment makes it the preferred group of compounds over esters for self-assembling amphiphiles. The slow hydrolysis of amides is the basis for the sustained release of guest molecules from the self-assembled micelles.

Materials and experimental procedures

Materials

The purification of solvents was done by the conventional distillation method. Anhydrous and deuterated solvents for NMR spectra were utilized as supplied by the local supplier. Synthetic compounds (purity): triglycerol (> 60%), *p*-toluenesulfonic acid (> 98%), 2, 2-dimethoxy propane (> 98%), triethylamine (> 99.5%), methyl dichloride (> 99%), oleoyl chloride (> 99%), stearoyl chloride (> 97%) and Dowex-50, were purchased from Sigma–Aldrich. The details of characterization techniques and analysis are given in supplementary information.

Experimental procedures

Preparation of TLC stains

1.0 g of $KMnO_4$ was dissolved in 150 mL water followed by the addition of 6.5 g K_2CO_3 and 1.0 mL 10% NaOH. This $KMnO_4$ stain is useful to visualize the oxidizable functional groups [amines, alcohols] after immersion into the stain followed by gentle heating.

General procedure for synthesis of G.n-amide compounds

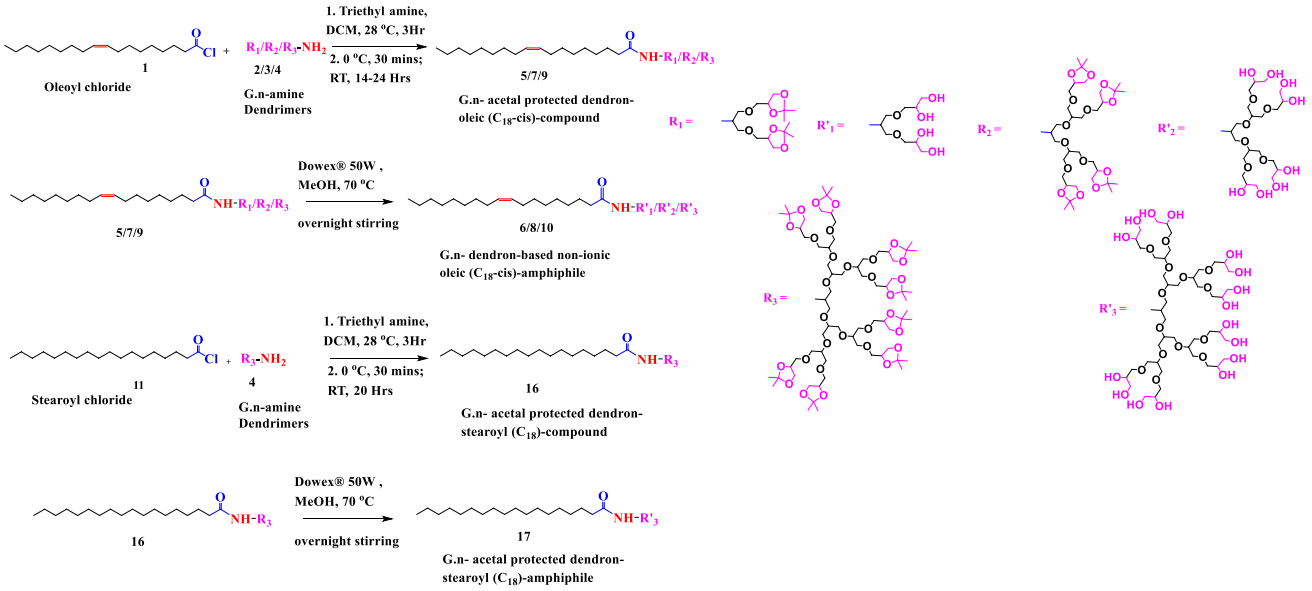
The synthesis of G.n-OH and G.n-NH2 dendrimer was carried out according to the published methodology [34, 35]. Triethylamine was added to a solution of the compound G.n-NH2 in dichloromethane, and the mixture was stirred continuously at 28 °C for 3–4 h. The addition of oleoyl chloride or stearoyl chloride was carried out carefully and slowly, the reaction conditions being kept at 0 °C. After the addition was complete, stirring was continued for 14–24 h at room temperature and the progress of the reaction was monitored by TLC. Once the reaction was complete, the unreacted acid was neutralized with sodium bicarbonate solution and extracted with DCM (350 mL). Anhydrous sodium sulfate was used to dry the extracted organic layer. To obtain the crude compounds, the solvent was evaporated on a vacuum rotary evaporator. The required products in pure form were obtained by column chromatography using ethyl acetate/hexane with different mixing ratios in order to obtain the various acetyl-protected intermediates 6, 8, 10 and 13, 15, 17 (Scheme 1) as a light yellow viscous oil.

N-(1,3-bis((2,2-dimethyl-1,3-dioxolan-4-yl) methoxy)propan-2-yl)oleamide [5]

R_f value—0.58 [n-hexane/ethyl acetate (7:3)]; yield: 80%; FTIR (cm^{-1}): 3380, 3179, 1650, 1622, 1591, 725, 696; ^1H NMR (400 MHz, $CDCl_3$, ppm): 7.62 (NH, m, 1H), 5.32 (-CH = CH-, m, 2H), 5.12–5.10 (-CH, m, 1H), 4.22–4.19 ($-CH_2$, m, 2H), 4.03–3.99 ($-CH_2$, m, 2H), 3.72–3.42

Table 1 Comparative physicochemical properties of dendron-based non-ionic amphiphiles with different linkers

Physicochemical properties	G.1-(C_{18}-cis) amphiphile		G.2-(C_{18}-cis) amphiphile		G.3-(C_{18}-cis) amphiphile	
	Ester liner	Amide linker	Ester liner	Amide linker	Ester liner	Amide linker
Particle size (nm)	157.1	93.00	93.00	55.00	111.10	13.00
CMC (mol/L × 10^{-5})	3.10	1.81	1.30	1.12	0.15	0.09
PDI value	0.144	0.132	0.369	0.248	0.195	0.189



Scheme 1 General synthesis route of dendritic amphiphiles with amide linker

(4 × -CH_2 and 2 × -CH, m, 10H), 2.29–2.25 (-CH_2, t, 2H, J = 7.6 Hz), 2.29–1.98 (3 × -CH_2, m, 6H), 1.38–1.32 (4 × -CH_3, 2 m, 12H); ^{13}C NMR (100.5 MHz, $CDCl_3$, ppm): 173.2, 130.4, 129.2, 109.1, 73.8, 72.0, 70.9, 70.2, 65.8, 34.4, 32.3, 29.1, 27.1, 26.2, 25.3, 23.9, 22.0, 14.1. Elemental analysis calculated. (%) for $C_{33}H_{61}NO_7$ (calculated mass 583.4448): C—67.89, H—10.53 and N—2.40; found: C—67.85, H—10.26 and N—2.58.

N-(1,13-bis(2,2-dimethyl-1,3-dioxolan-4-yl)-4,10-bis(((2,2-dimethyl-1,3-dioxolan-4-yl) methoxy) methyl)-2,5,9,12-tetraoxatridecan-7-yl) oleamide [7]

R_f value—0.52 [n-hexane/ethyl acetate (7:3)]; yield: 68%; FTIR (cm^{-1}): 3375, 3189, 1645, 1576, 1546, 758, 678; ^{1}H NMR (400 MHz, $CDCl_3$, ppm): 7.90 (-NH, m, 1H), 5.25–5.22 (-CH = CH-, m, 2H), 5.14–5.12 (-CH, m, 1H), 4.13–4.11 (2 × -CH_2, m, 4H), 3.94–3.90 (2 × -CH_2, m, 4H). 3.63–3.59 (4 × -CH_2, t, 8H, J = 7.2 Hz), 3.45–3.37 (6 × -CH_2 and 6 × -CH, m, 18H), 2.22–2.18 (-CH_2, t, 2H, J = 7.2 Hz), 1.92–1.87 (2 × -CH_2, m, 4H), 1.51–1.48 (-CH_2, m, 2H), 1.29–1.14 (8 × -CH_3, m, 24H), 1.12–1.08 (10 × -CH_2, m, 20H), 0.75 (-CH_3, t, 3H, J = 5.6 Hz); ^{13}C NMR spectrum (100.5 MHz, $CDCl_3$): δ = 173.0, 129.2, 108.1, 78.3, 76.8, 74.1, 72.1, 71.0, 70.2, 70.1, 69.1, 66.2, 34.1, 32.1, 29.9, 26.2, 25.2, 22.1, 14.0. Elemental analysis calculated. (%) for $C_{51}H_{93}NO_{15}$ (calculated mass 959.6545): C—63.79, H—09.76 and N—1.46; found: C—63.85, H—09.78 and N—01.48.

N-(7,13-bis(((1,3-bis((2,2-dimethyl-1,3-dioxolan-4-yl) methoxy)propan-2-yl)oxy)methyl)-1,19-bis(2,2-dimethyl-1,3-dioxolan-4-yl)-4,16-bis(((2,2-dimethyl-1,3-dioxolan-4-yl)methoxy)methyl)-2,5,8,12,15,18-hexaoxanonadecan-10-yl)oleamide [9]

R_f value—0.50 [n-hexane/ethyl acetate (7:3)]; yield: 57%; FTIR (cm^{-1}): 3360, 2931, 2848, 1739, 1459, 1019, 725, 693; ^{1}H NMR (400 MHz, $CDCl_3$, ppm): 7.50 (-NH, m, 1H), 5.27–5.25 (-CH = CH-, m, 2H), 5.02–4.97 (-CH, m, 1H), 4.17–4.5 (4 × -CH_2, m, 8H), 4.02–3.97 (4 × -CH_2, t, 8H, J = 7.2 Hz), 3.63–3.52 (2 × -CH_2 and 6 × -CH, m, 10H), 3.48–3.45 (20 × -CH_2 and 8 × -CH, m, 48H), 2.24–2.20 (-CH_2, t, 2H, J = 7.2 Hz), 1.98–1.82 (2 × -CH_2, m, 4H), 1.54–1.48 (-CH_2, m, 2H), 1.34–1.28 (16 × -CH_3, m, 48H), 1.19–1.10 (10 × -CH_2, m, 20H), 0.81 (-CH_3, t, 3H, J = 5.6 Hz); ^{13}C NMR spectrum (100.5 MHz, $CDCl_3$): δ = 174.1, 137.1, 109.2, 77.3, 76.5, 71.1, 32.0, 29.7, 29.3, 23.4, 14.5. Elemental analysis calculated. (%) for $C_{87}H_{157}NO_{31}$ (calculated mass 1712.0740): C—60.96, H—09.17 and N—00.82; found: C—60.85, H—09.26 and N—00.81.

N-(1,3-bis((2,2-dimethyl-1,3-dioxolan-4-yl) methoxy)propan-2-yl)stearamide [12]

R_f value—0.64 [n-hexane/ethyl acetate (7:3)]; yield: 85%; FTIR (cm^{-1}): 3375, 3185, 1630, 1591, 1564, 754, 665; ^{1}H NMR (400 MHz, $CDCl_3$, ppm):7.62 (-NH, m, 1H), 5.12–5.10 (-CH, m, 1H), 4.22–4.19 (-CH_2, m, 2H), 4.03–3.98 (-CH_2, m, 2H), 3.72–3.42 (4 × -CH_2 and 2 × -CH,

m, 10H), 2.29–2.00 (2×-CH_2, m, 4H), 1.38–1.33 (4×-CH_3, 2 m, 12H), 1.28–1.25 (14×-CH_2, m, 28H), 0.85 (-CH_3, t, 3H, J = 6.8); ^{13}C NMR (100.5 MHz, $CDCl_3$, ppm): 173.3, 130.0, 129.5, 109.1, 74 0.1, 72.2, 71.0, 69.8, 66.4, 34.1, 32.3, 28.9, 27.5, 26.3, 25.6, 24.1, 21.7, 14.1. Elemental analysis calculated. (%) for $C_{33}H_{63}NO_7$ (calculated mass 585.4605): C—67.65, H—10.84 and N—2.39; found: C—67.85, H—10.86 and N—2.40.

N-(1,13-bis(2,2-dimethyl-1,3-dioxolan-4-yl)-4,10-bis(((2,2-dimethyl-1,3-dioxolan-4-yl) methoxy) methyl)-2,5,9,12-tetraoxatridecan-7-yl) stearamide [14]

R_f value—0.55 [n-hexane/ethyl acetate (7:3)]; yield: 70%; FTIR (cm^{-1}): 3375, 3185, 1630, 1591, 1564, 754, 665; ^{1}H NMR (400 MHz, $CDCl_3$, ppm): 7.50 (-NH, m, 1H), 5.12–5.10 (-CH, m, 1H), 4.22–4.19 (2×-CH_2, m, 4H), 4.03–3.99 (2×-CH_2, m, 4H), 3.72–3.60 (6×-CH_2 and 6×-CH, m, 18H), 3.52–3.45 (3×-CH_2, m, 6H), 2.29–1.98 (2×-CH_2, m, 4H), 1.59 (-CH_2, t, 2H, J = 7.2 Hz), 1.29–1.14 (8×-CH_3, m, 24H), 1.28–1.24 (14×-CH_2, m, 28H), 0.85 (-CH_3, t, 3H, J = 6.8 Hz); ^{13}C NMR (100.5 MHz, $CDCl_3$, ppm): 173.2, 129.8, 129.2, 109.4, 73.9, 72.3, 71.1, 70.0, 66.3, 34.1, 32.3, 29.0, 27.1, 26.2, 25.4, 24.1, 22.2, 14.1. Elemental analysis calculated. (%) for $C_{51}H_{95}NO_{15}$ (calculated mass 961.6702): C—63.65, H—09.95 and N—1.46; found: C—63.85, H—09.96 and N—1.58.

N-(7,13-bis(((1,3-bis((2,2-dimethyl-1,3-dioxolan-4-yl) methoxy)propan-2-yl)oxy)methyl)-1,19-bis(2,2-dimethyl-1,3-dioxolan-4-yl)-4,16-bis(((2,2-dimethyl-1,3-dioxolan-4-yl)methoxy)methyl)-2,5,8,12,15,18-hexaoxanonadecan-10-yl)stearamide [16]

R_f value—0.49 [n-hexane/ethyl acetate (7:3)]; yield: 85%; FTIR (cm^{-1}): 3365, 2951, 2850, 1739, 1462, 1020, 735, 683; ^{1}H NMR (400 MHz, $CDCl_3$, ppm): 7.48 (-NH, m, 1H), 5.03–5.00 (-CH, m, 1H), 4.17–4.15 (4×-CH_2, m, 8H), 4.02–3.97 (4×-CH_2, t, 8H, J = 7.2 Hz), 3.61–3.54 (2×-CH_2 and 6×-CH, m, 10H), 3.52–3.47 (20×-CH_2 and 8×-CH, m, 48H), 2.24–2.20 (-CH_2, t, 2H, J = 7.2 Hz), 1.57–1.52 (-CH_2, m, 2H), 1.35–1.24 (16×-CH_3, m, 48H), 1.11–1.19 (14×-CH_2, m, 28H), 0.88 (-CH_3, t, 3H, J = 5.6 Hz); ^{13}C NMR (100.5 MHz, $CDCl_3$): δ = 174.5, 137.4, 109.1, 77.4, 76.0, 71.1, 32.0, 30.1, 29.5, 23.3, 14.2. Elemental analysis calculated (%) for $C_{87}H_{159}NO_{31}$ (calculated mass 1714.0896): C—60.92, H—09.34 and N—00.82; found: C—60.85, H—09.26 and N—00.80.

General procedure for G.n-amide amphiphiles synthesis (deprotection of G.n-amide)

The ion-exchange resin Dowex® 50 W was used to break the acetyl-protected groups. The protected compound was taken in methanol (25 mL) and was refluxed overnight at 70 °C. The resin was filtered at room temperature. The filtrate material was concentrated under vacuum to get G.n-amide amphiphilic compounds 6, 8, 10 and 13, 15, 17 (Scheme 1) as products.

1,3-Bis(2,3-dihydroxypropoxy)propan-2-yl oleate [6]

R_f value—0.48 [chloroform/methanol (9:1)]; yield: 87%; FTIR (cm^{-1}): 3582, 3079, 1616, 1586, 1577, 720 and 689; ^{1}H NMR (400 MHz, CD_3OD, ppm): 5.32 (-CH = CH-, m, 2H, J = 6.0 Hz), 5.12 (-CH, m, 1H), 3.76–3.58 (6×-CH_2 and 2H, m, 2×-CH, m, 14H), 2.05–2.02 (-CH_2, t, 2H, J = 7.6 Hz), 2.01–1.98 (2×-CH_2, q, 4H, J = 6.8), 1.62–1.59 (-CH_2, t, 2H, J = 7.2), 1.33–1.29 (10×-CH_2, d, 20H), 0.91 (-CH_3, t, 3H, J = 6.8 Hz); ^{13}C NMR (100.5 MHz, CD_3OD): 173.92, 129.46, 79.99, 70.83, 63.05, 29.01, 28.87, 26.80, 13.15. Elemental analysis calculated. (%) for $C_{27}H_{53}NO_7$ (calculated mass 503.3822): C—64.38, H—10.61 and N—02.78; found: C—64.35, H—10.56 and N—02.80.

6,12-Bis((2,3-dihydroxypropoxy) methyl)-1,2,16,17-tetrahydroxy-4,7,11,14-tetraoxaheptadecan-9-yl oleate [8]

R_f value—0.40 [chloroform/methanol (9:1)]; yield: 90%; FTIR (cm^{-1}): 3370, 3025, 2856, 1732, 1460, 1019, 728, 690; ^{1}H NMR (400 MHz, CD_3OD, ppm): 5.33 (-CH = CH-, m, 2H), 5.09–5.08 (-CH, m, 1H), 3.55–3.43 (15× –CH_2 and 4×-CH, m, 34H), 2.35–2.31 (-CH_2, t, 2H, J = 6.8 Hz), 2.02–1.99 (2×-CH_2, m, 4H), 1.61–1.56 (-CH_2, m, 2H), 1.31–1.27 (10×-CH_2, d, 20H), 0.88 (-CH_3, t, 3H, J = 7.2 Hz); ^{13}C NMR spectrum (100.5 MHz, CD_3OD): δ = 172.81, 129.41, 78.31, 77.00, 74.36, 72.22, 71.34, 71.14, 69.78, 68.66, 66.47, 34.08, 31.63, 29.04, 26.52, 25.15, 22.41, 13.88. Elemental analysis (%) for $C_{39}H_{77}O_{15}$ (calculated mass 800.0264): C—58.55, H—09.70 and N—01.75; found: C—58.57, H—09.66 and N—01.77.

9,15-Bis(((1,3-bis(2,3-dihydroxypropoxy) propan-2-yl)oxy) methyl)-6,18-bis((2,3-dihydroxypropoxy) methyl)-1,2,22,23-tetrahydroxy-4,7,10,14, 17,20-hexaoxatricosan-12-yl oleate [10]

R_f value—0.32 [chloroform/methanol (9:1)]; yield 81%; FTIR (cm^{-1}): 3375, 2932, 2860, 1740, 1465, 1023, 725, 702. ^{1}H NMR (400 MHz, CD_3OD, ppm): 5.31 (-CH = CH-, m, 2H), 5.11–5.09 (-CH, m, 1H), 3.86–3.22 (30 × -CH_2 and 14 × -CH, m, 74H), 2.32–2.27 (-CH_2, t, 2H, J = 6.8 Hz), 2.14–2.00 (2 × -CH_2, m, 4H), 1.68–1.57 (-CH_2, m, 2H), 1.28–1.26 (10 × -CH_2, d, 20H), 0.87 (-CH_3, t, 3H, J = 7.2 Hz); ^{13}C NMR (100.5 MHz, CD_3OD): δ = 174.52, 129.46, 129.58, 78.50, 72.65, 71.63, 70.87, 69.82, 63.13, 47.93, 47.29, 29.54, 28.89, 26.85, 13.21. Elemental analysis (%) for $C_{63}H_{125}O_{31}$ (calculated mass 1392.6547): C—54.33, H—09.05 and N—01.01; found: C—54.29, H—08.98 and N—00.98.

N-(1,3-bis(2,3-dihydroxypropoxy) propan-2-yl)stearamide [13]

R_f value—0.51 [chloroform/methanol (9:1)]; yield: 85%; FTIR (cm^{-1}): 3582, 3079, 1616, 1586, 1577, 720 and 689. ^{1}H NMR (400 MHz, CD_3OD, ppm): 5.09 (-CH, m, 1H), 4.83–4.80 (2 × -CH_2, m, 4H), 3.86–3.27 (4 × -CH_2, m, 8H), 2.32–2.27 (-CH_2, t, 2H, J = 7.6 Hz), 2.14–2.00 (2 × -CH_2, q, 2H, J = 6.8 Hz), 1.68–1.57 (-CH_2, t, 2H, J = 7.2 Hz), 1.28–1.26 (10 × -CH_2, d, 26H), 0.87 (-CH_3, t, 3H, J = 6.8 Hz); ^{13}C NMR (100.5 MHz, CD_3OD): 173.12, 74.53, 72.18, 70.84, 69.95, 69.59, 34.31, 31.84, 29.25, 27.11, 26.67, 25.35, 24.89, 22.62, 14.06. Elemental analysis calculated. (%) for $C_{27}H_{55}NO_7$ (calculated mass 505.3979): C—64.12, H—10.96 and N—02.77; found: C—64.15, H—10.87 and N—02.80.

N-(6,12-bis((2,3-dihydroxypropoxy) methyl)-1,2,16,17-tetrahydroxy-4,7,11,14-tetraoxaheptadecan-9-yl)stearamide [15]

R_f value—0.42 [chloroform/methanol (9:1)]; yield: 86%; FTIR (cm^{-1}): 3370, 3025, 2856, 1732, 1460, 1019, 728, 690. ^{1}H NMR (400 MHz, CD_3OD, ppm): 5.09 (-CH, m, 1H), 4.12–3.33 (15 × -CH_2 and 4 × -CH, m, 34H), 2.02–2.01 (2 × -CH, m, 2H), 1.85–1.75 (-CH_2, t, 2H, J = 6.8 Hz), 1.54–1.36 (-CH_2, m, 2H), 1.27–1.10 (10 × -CH_2, d, 26H), 0.75 (-CH_3, t, 3H, J = 7.2 Hz); ^{13}C NMR spectrum (100.5 MHz, CD_3OD): δ = 173.20, 78.61, 77.47, 77.16, 76.85, 74.64, 72.52, 71.62, 71.42, 66.91, 66.77, 31.92, 29.7729.16, 27.23, 26.79, 25.43, 22.69, 14.14. Elemental analysis (%) for $C_{39}H_{79}O_{15}$ (calculated mass 802.0423): C—58.40, H—09.93 and N—01.75; found: C—58.45, H—09.66 and N—01.82.

N-(9,15-bis(((1,3-bis(2,3-dihydroxypropoxy)propan-2-yl)oxy)methyl)-6,18-bis((2,3-dihydroxypropoxy) methyl)-1,2,22,23-tetrahydroxy-4,7,10,14,17,20-hexaoxatricosan-12-yl)stearamide [17]

R_f value—0.38 [chloroform/methanol (9:1)]; yield: 78%; FTIR (cm^{-1}): 3375, 2975, 2860, 1740, 1470, 1023, 725, 702. ^{1}H NMR (400 MHz, CD_3OD, ppm): 5.02 (-CH, m, 1H), 4.16–3.10 (31 × -CH_2 and 14 × -CH, m, 76H), 1.95–1.82 (-CH_2, t, 2H, J = 6.8 Hz), 1.34–1.28 (-CH_2, m, 2H), 1.19–1.10 (14 × -CH_2, d, 26H), 0.82 (-CH_3, t, 3H, J = 7.2 Hz); ^{13}C NMR spectrum (100.5 MHz, CD_3OD): δ = 173.20, 78.53, 70.94, 63.15, 29.58, 29.09, 28.97, 13.27. Elemental analysis (%) for $C_{63}H_{127}O_{31}$ (calculated mass 1394.6706): C—54.25, H—09.18 and N—01.00; found: C—54.09, H—09.08 and N—01.06.

Measurements

FTIR spectroscopy

The FTIR spectra were recorded at PerkinElmer's Spectrum Two Nutraceutical & TCM System Part Number: L160000W by using PerkinElmer Spectrum, version 10.5.3. Dehydrated potassium iodide was used for pallet making.

NMR spectroscopy

The ^{1}H and ^{13}C spectra of the synthesized compounds were recorded using the deuterated methanol and chloroform on a Bruker DRX-400, 400 MHz, spectrometer. Deuterated chloroform ($CDCl_3$) and deuterated methanol (CD_3OD) were used as solvents.

Dynamic light scattering (DLS)

DLS was carried out at room temperature using quartz cuvettes with 10 mm path length to determine the polydispersity index and the average size (the results are calculated automatically by the software and the number of runs per sample was 50) of the scattered particles by Zeta-Seizer Nano-S of Malvern instruments. Five milligrams of amphiphile sample in 10 mL DI water was stirred overnight at room temperature, and after stabilization of the micelle, it was filtered with a 0.2-m nylon filter for DLS studies.

Cryo-transmission electron microscopy

Cryo-TEM of frozen aqueous samples was carried out with an FEI Tecnai TF20. A drop of an aqueous solution of the self-assembled aggregate sample was deposited on a copper grid (coated with carbon). A Vitrobot (Mark IV) automated device was used to vitrify aqueous samples for cryo-TEM. The samples were analyzed by TEM with a FEI Tecnai TF20 microscope at an accelerating voltage of 200 kV.

Surface tension and critical micelle concentration (CMC)

The hanging drop method was carried out using the drop shape analyzer DSA25 from Kruss GmbH with contact edge estimation range of 0°–180° (resolution: 1 0.1°) and surface pressure ranges from 0.01–1000 Nm/m (resolution: 0.1 Nm/m) used. The main investigation was assessed by PerkinElmer 2400. The surface tension of eight samples at different concentrations was observed, and the graphs were plotted against the respective concentration to determine the CMC.

Surface excess concentration (Γ)

The area-related concentration of a non-ionic amphiphile at the surface or interface is known as surface excess concentration. Mathematically, it can calculate using Eq. (1):

$$\Gamma = \frac{-1}{RT}\,\frac{d\gamma}{d\ \ln\ C} \tag{1}$$

where R is the universal gas constant, T is the absolute temperature and C is the concentration.

Molecular occupied area (A)

Molecular occupied area by a non-ionic amphiphilic molecule at the solution–air interface was calculated from the CMC curve using Eq. (2):

$$A = \frac{1 \times 10^{18}}{\Gamma N} \tag{2}$$

where Γ is the surface excess concentration and N is the Avogadro's number (6.023×10^{23}).

Results and discussion

Synthesis

The synthesis of G.n-NH_2 dendrons was completed according to the reported conventional methodology [35]. The conversion of the amine–polyglycerol dendrimer to the amide amphiphiles (Scheme 1 and Fig. 1) was confirmed by ^{1}H-NMR, ^{13}C-NMR and IR spectroscopic data and comparing it with the reported data [35, 36]. The methodology was also used to characterize the synthesized saturated and unsaturated non-ionic amphiphile with different generations of dendrimers. ^{13}C Spectra of compound 5 showing a peak at 109.34 ppm signify –C=C–, which is absent in spectra of compound 12. The sharp peaks at about 1650 cm^{-1} and 3380 cm^{-1} in IR spectra of compounds confirm the presence

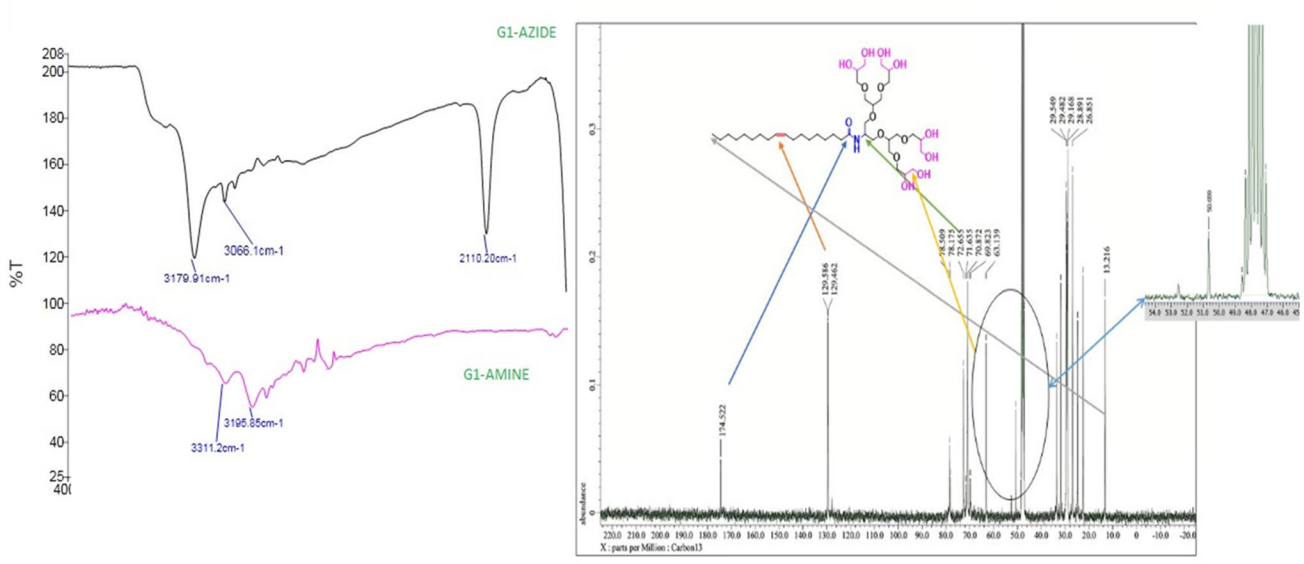


Fig. 1 **A** IR spectra of dendritic compounds, with azide and amine core, **B** ^{13}C spectra of compound 8

of amide group in the compound, while the broad peak at 3582 cm^{-1} in the spectra is because of free hydroxyl groups.

Average particle size and morphological studies

The DLS investigations on self-aggregated particles of amphiphiles 6, 8, 10 and 13, 15, 17 showed that the shape and size of micelles formed by the dendritic amphiphile depend on the structural character of the dendritic unit [8, 21]. The amphiphiles can frame an aggregation with a unimodal character in the nano-size range. The investigations showed that the particle size of amphiphiles changes with the change in the hydrophilic and hydrophobic parts. The unsaturation in the hydrophobic chain reacts by increasing the size of the particles. The smallest particles are formed by an amide-bonded G.3 dendron-based non-ionic (C_{18}) amphiphile 17, i.e., 7.5 nm, while the largest 93.0 nm are formed by an amide-bonded G.1 dendron-based non-ionic (C_{18}-cis) amphiphile 6 (Table 2). The micelle accumulations have essentially low polydispersity index (PDI) qualities and provide information about the size of the micelle shape (Table 2). The SEM and cryo-TEM studies were performed on a CMC above (Fig. 2 and Fig. 3). The self-aggregating micelles are framed due to the structure-specific interaction of amphiphiles with water [11, 37]. It has been shown that the aggregation number decreases with the enlargement of the hydrophilic head bundle zone of amphiphiles. This leads to a further decrease in size across the micelle (Fig. 2). The entanglement of the aggregation threads is high in the case of amphiphiles based on amphiphiles based on G.1 dendrons (C_{18}-cis) ionic amphiphiles 8, 10 and 15, 17 (Figs. 3 and 4). Figure 4 shows a schematic representation of the self-aggregation of various dendritic non-ionic amphiphiles in water.

Table 2 Various physicochemical and adsorption properties of non-ionic amide amphiphiles and sustained dye released percentage of encapsulation efficiency of amide-linked dendron-based amphiphiles

Amide-linked dendritic amphiphile	Comp. no.	Particle size (nm)	PDI	Aggregation morphology	Yield (%)
G.1-(C-18-cis) non-ionic amphiphile	6	93.00	0.132	Rod Shape	85
G.2-(C-18-cis) non-ionic amphiphile	8	55.00	0.248	Spherical	86
G.3-(C-18-cis) non-ionic amphiphile	10	13.00	0.189	Spherical	78
G.1-(C-18) non-ionic amphiphile	13	28.00	0.124	Rod Shape	87
G.2-(C-18) non-ionic amphiphile	15	11.00	0.193	Spherical	90
G.3-(C-18) non-ionic amphiphile	17	07.50	0.112	Spherical	81

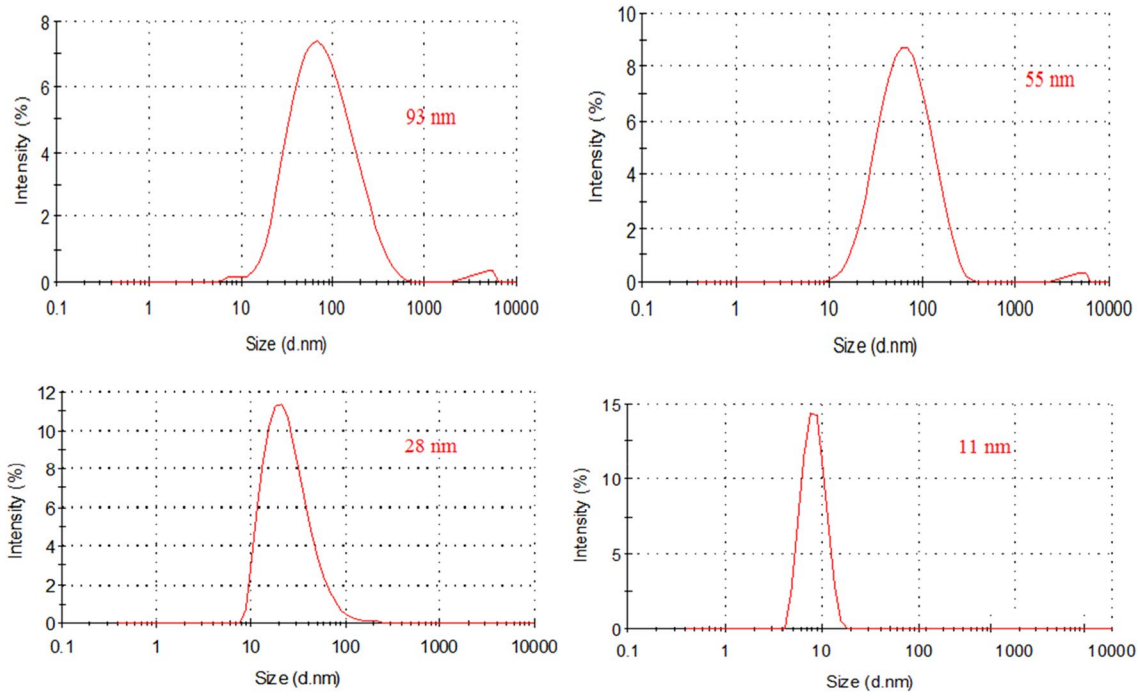


Fig. 2 Average particle size determination of synthesized amphiphiles by DLS. First row: L–R: compounds **6**, **8**. Second row L–R: compounds 13, 15

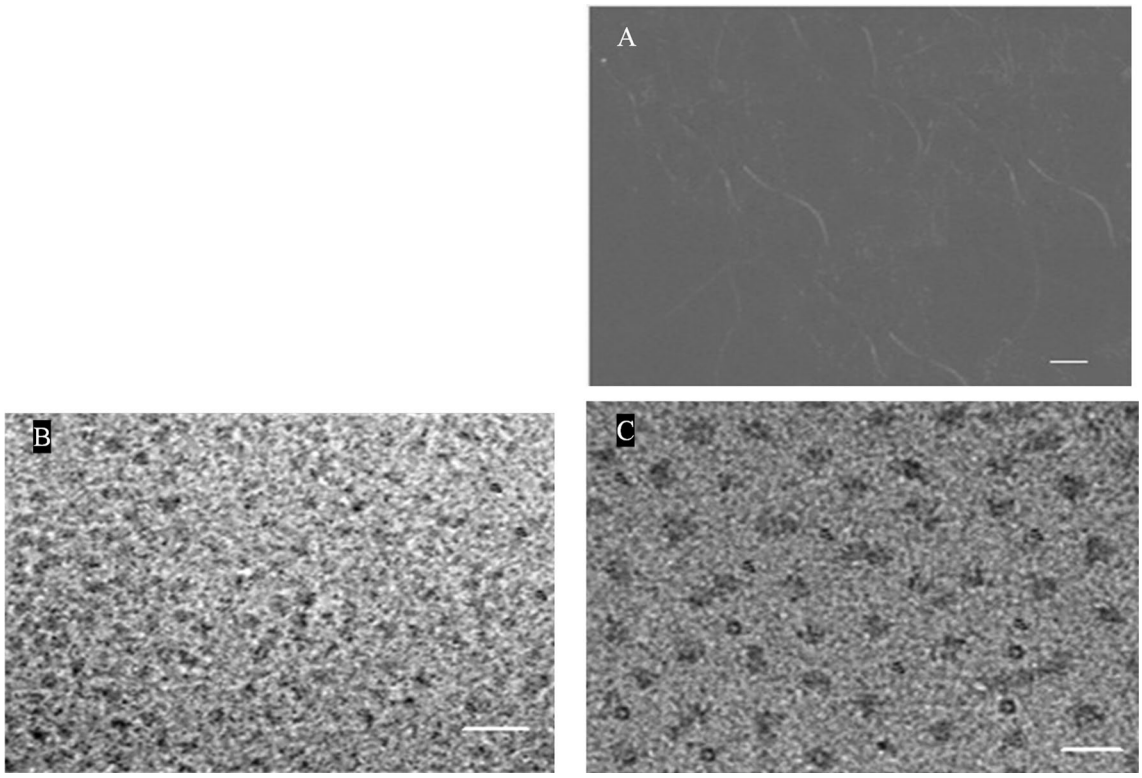


Fig. 3 A: SEM image of compound 8 *(scale 200 nm)*; **B** cryo-TEM pictures of watery self-assembled structural arrangement of amphiphile 13; **C** cryo-TEM pictures of watery self-assembled structural arrangement of amphiphile 15 (*scale: 100 nm*)

Critical micelle concentration (CMC) and adsorption properties

All synthesized non-ionic amphiphiles on amphiphiles that are linked with amide have good solubility in neutral water at room temperature. The lowering of the Krafft temperature of the amphiphiles was attributed to two factors independent of a long hydrocarbon chain length. First, the cis double bond in the C-9 position causes the unsaturated hydrophobic tails to shrink [38]. Second, an increase in the hydrophilic head groups increases the solubility of amphiphiles in water [39, 40]. It has also been observed that these long-chain dendritic amphiphiles immediately self-organize into viscoelastic worm-like or spherical micelles. The relationship between amphiphile concentration and surface tension was checked using the pendant drop technique [41] (Fig. 5). The influence of the PG dendron generations G.1–G.3 as increasing amphiphile units and saturation of the aliphatic chain was investigated. The results show that the representation of the unsaturation in the hydrocarbon chain can be viewed as an increase in the hydrophobicity of the hydrocarbon chain since the delocalized electrons increase the polarizability of the chain. In addition, the unsaturation reduces the adaptability of the aliphatic chain. At the point where there is a carbon–carbon double bond in the hydrocarbon chain, the CMC increases. Bigger size of hydrophilic head and saturation in hydrophobic tail prompts a lower CMC. A larger size of the hydrophilic head and saturation in the hydrophobic tail result in a lower CMC. The CMC values obtained for the amphiphiles in relation to the dendritic generations (G.1, G.2, G.3) are 1.81×10^{-5} mol L^{-1}, 1.12×10^{-5} mol L^{-1} and 9.0×10^{-7} mol L^{-1} for (C_{18}-cis) amphiphiles 6, 8, 10 (Fig. 5), whereas 1.0×10^{-6} mol L^{-1}, 8.0×10^{-7} mol L^{-1} and 6.0×10^{-7} mol L^{-1} for (C_{18}) amphiphiles are 13, 15 and 17 (Fig. 5 and Table 3), respectively. The lower CMC values of compounds 13, 15, 17 compared to compounds 6, 8, 10 are driven by the strong hydrophobic interaction. Both amphiphiles 10 and 17 have a G.3 dendritic head that has a hydrogen-bond interaction with water, but the olefinic chain in compound 10 has a distributed polarity because the C_9–C_{10} double bond is in the hydrophobic tail. This more polar nature of the hydrophobic tail in (C_{18}-cis) amphiphiles than in (C_{18}) amphiphiles

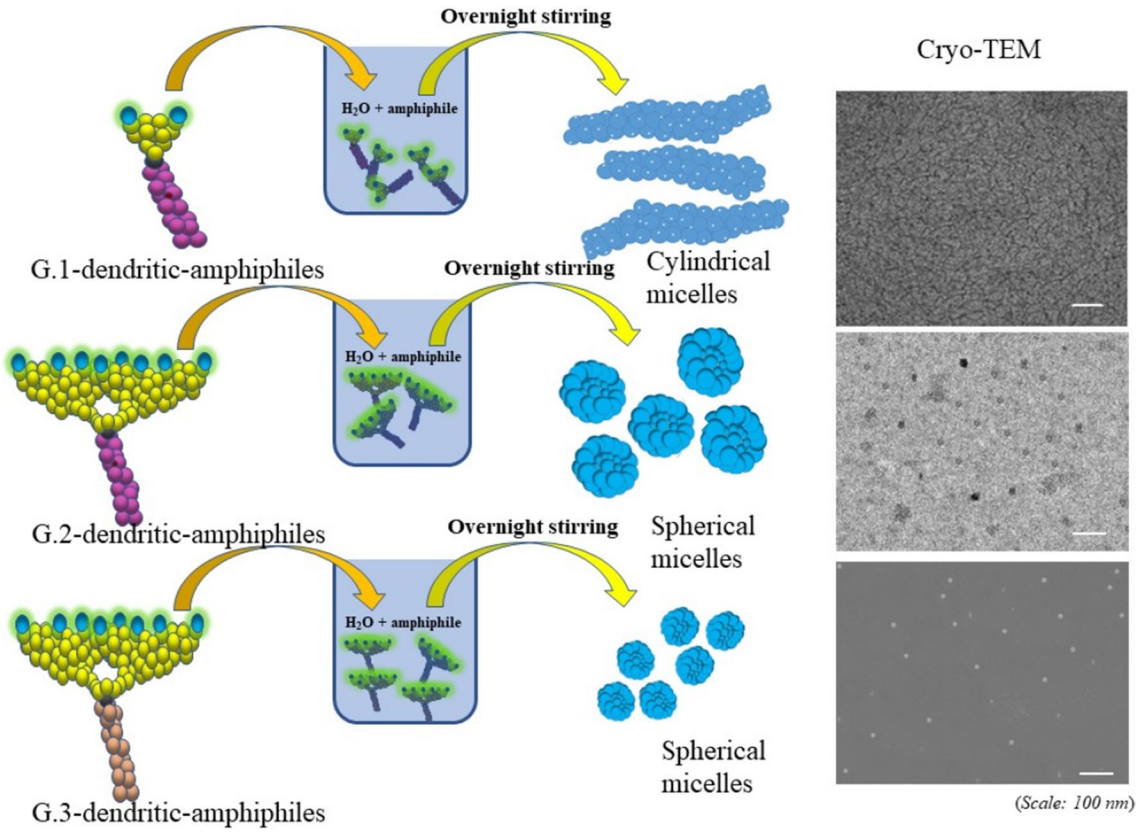


Fig. 4 Schematic of self-aggregation and cryo-TEM images of different dendritic non-ionic amphiphiles in water. The worm-like or thread-like morphology is observed for G.1-dendron-based non-ionic amphiphiles and spherical self-aggregated micelles for G.2- and G.3-dendron-based non-ionic amphiphiles [top-down: compounds 6, 10, 17]. With the same generation dendrimers, amphiphiles have same morphology but different particle sizes, which was confirmed by DLS and cryo-TEM studies (*scale: 100 nm*)

reacted with a higher CMC value. Previous publications do not report the decrease in CMC due to the increase in head size and the increase in the hydrophilic nature of the head group [8, 17, 41]. Various interface parameters of amphiphiles in an aqueous medium can also be calculated from the CMC curve. Gibbs isothermal equations can be used to calculate the adsorption properties of the aqueous solutions of the amphiphiles. In the case of (C_{18}) amphiphiles, the surface excess concentration [42] decreased with increasing dendron formation of the dendritic unit, while no such specific trend was observed for (C_{18}-cis) amphiphiles. The surface tension of amphiphiles on CMC increases with the increase in the generation of dendrons from G.1 to G.3 [41, 43]. Trends given in Table 1 for the G.1 dendron having CMC surface tension within the range of 42–52 are in accordance with the literature [44].

Conclusions

A family of amide-linked dendron-based non-ionic (C_{18}) and (C_{18}-*cis*) amphiphiles have been synthesized utilizing biocompatible synthon as starting materials. The amphiphiles get self-assembled to frame supramolecular micellar structures of spherical or rod-like morphology in water. The variation in micellar morphology formed depends upon the ratio of size of the hydrophilic head to hydrophobic chain of the amphiphiles. The presence of unsaturation in the hydrophobic chain demonstrated the lowering of CMC value. The DLS and Cryo-TEM studies demonstrated the nano-range of self-assembled aggregates. Significant values of adsorption parameters have also been observed. Amide-linked dendron-based amphiphile has the potential to serve as an intracellular delivery vehicle. The integrated amphiphiles are biocompatible and provide future scopes.

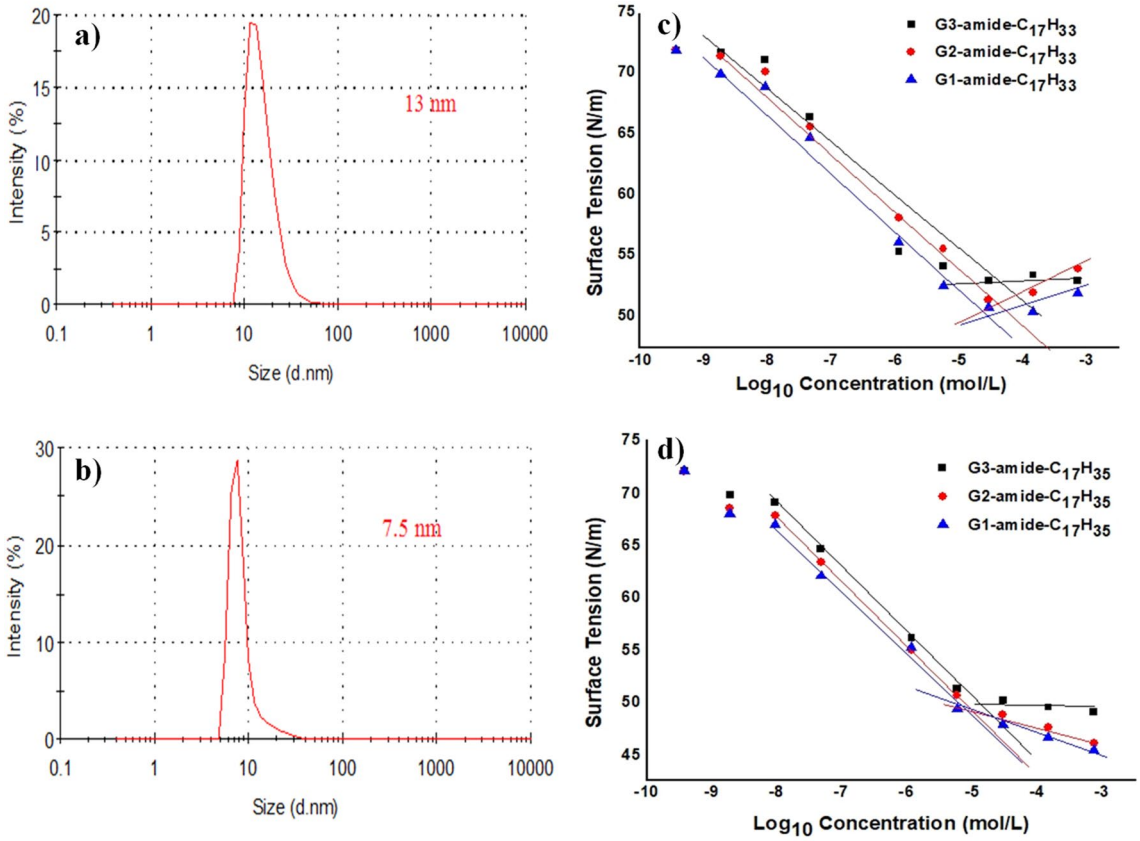


Fig. 5 (**a**) and (**b**) Average particle size determination of synthesized amphiphiles by DLS: compound **10** and compound 17, respectively. Graph (**c**) concentration of amphiphiles 6, 8, 10 [amide-linked G.n-dendron-based non-ionic (C_{18}) amphiphile]. Graph (**d**) concentration of amphiphiles 13, 15, 17 [amide-linked G.n-dendron-based non-ionic (C_{18}-cis) amphiphile] against surface tension

Table 3 Various physicochemical and adsorption properties of non-ionic amide amphiphiles and sustained dye released percentage of encapsulation efficiency of amide-linked dendron-based amphiphiles

Amide-linked dendritic amphiphile	Comp. no.	CMC (mol/L × 10^{-5})	$10^6\ \Gamma_{max}$ (mol m^{-2})	A	γ_{CMC} (mN/m)
G.1-(C_{18}-cis) non-ionic amphiphile	6	1.81	3.49	0.475	50.18
G.2-(C_{18}-cis) non-ionic amphiphile	8	1.12	3.27	0.507	51.15
G.3-(C_{18}-cis) non-ionic amphiphile	10	0.09	3.10	0.535	52.88
G.1-(C_{18}) non-ionic amphiphile	13	0.10	3.39	0.489	49.74
G.2-(C_{18}) non-ionic amphiphile	15	0.08	2.93	0.566	49.85
G.3-(C_{18}) non-ionic amphiphile	17	0.06	2.72	0.610	49.96

Critical micelle concentration (CMC), surface tension at CMC (γ_{CMC}), surface excess concentration (Γ), area occupied by a molecule (A) (nm^2/molecule) at the air/solution interface of amphiphiles calculated at 28 $^{\circ}$C and neutral pH

Declarations

Conflicts of interest The authors declare that they have no conflict of interest.

References

1. D. Shi, D. Nguyen, M. Maaloum, J. Gallani, D. Felder-Flesch, M.P. Krafft, Molecules **24**, 4114–4127 (2019)
2. D.A. Tomalia, L.S. Nixon, D.M. Hedstrand, Biomolecules **10**, 642–698 (2020)
3. E. Fedeli, A. Lancelot, J.M. Dominguez, J.L. Serrano, P. Calvo, T. Sierra, Nanomaterials **9**, 161–175 (2019)
4. J. Shi, A.R. Votruba, O.C. Farokhzad, R. Langer, J. Shi, A.R. Votruba, O.C. Farokhzad, Nano Lett. **10**, 3223–3230 (2010)
5. H. Zhu, L. Shangguan, B. Shi, G. Yu, F. Haung, Mat. Chem. Front. **2**, 2152–2174 (2018)
6. N. Zacharopoulos, I.G. Economou, Macromolecules **35**, 1814–1821 (2002)

7. S. Senapati, A.K. Mahanta, S. Kumar, P. Maiti, Signal Transduction and Targeted. Therapy **3**, 7 (2018)
8. A. Kumar, S. Tyagi, R. Singh, Y.K. Tyagi, New J. Chem. **43**(2), 1025–1031 (2019)
9. Y. Ohta, Y. Abe, K. Hoka, E. Baba, Y.-P. Lee, C.-A. Dai, T. Yokozawa, Polym. Chem. **10**, 4246–4251 (2019)
10. P.K. Vuram, U. Subuddhi, S.T. Krishnaji, A. Chadha, A.K. Mishra, Eur. J. Org. Chem. **26**, 5030–5040 (2010)
11. Xi. Liu, I. Gitsov, Macromolecules. **52**, 5563–5573 (2019)
12. S. Tripathy, M.K. Das, JAPS. **3**(09), 142–149 (2013)
13. M. Gupta, V. Sharma, Res. J. Chem. Sci. **1**(2), 135 (2011)
14. S. Mirsharghi, K. D. Knudsen, S. Bagherifam, B. Nystro¨m, U. Boas, New. J. Chem. **40**, 3597–3611 (2016)
15. A. Popielec, E. Fenyvesi, K. Yannakopoulou, T. Loftsson, Pharmazie. **71**, 68–75 (2016)
16. B. Yoo, J.K. Shah, Y. Zhu, E.J. Maginn, Soft Matter **10**, 8641–8651 (2014)
17. N. Kalra, G. Jeyabalan, G. Singh, S. Choudhary, JIPBS. **3**(2), 193 (2016)
18. P. Degen, D.C.F. Wieland, C. Strötges, Langmuir **31**(43), 11851–11857 (2015)
19. Xi. Zhang, H. Xing, Y. Zhao, Z. Ma, Pharmaceutics. **10**, 74 (2018)
20. S. Kalepu, V. Nekkanti, Acta Pharm. Sin. B. **5**(5), 442–453 (2015)
21. I.O. Lebedeva, E.B. Zhulina, O.V. Borisov, Macromolecules **52**, 3655–3667 (2019)
22. Xi Zhang, and C. Wang, Chem. Soc. Rev. **40**, 94–101 (2011)
23. A. Banerjee, A. Tam, M. Dutt, Nanoscale Adv. **3**, 725–737 (2021)
24. M.S. Bakshi, A. Kaura, R.K. Mahajan, T. Yoshimura, K. Esumi, Colloids and Surfaces A: Physicochem. Eng. Aspects. **246**, 39–48 (2004)
25. B. Lin, A.V. McCormick, H.T. Davis, R. Strey, J. Colloid Interface Sci. **291**, 543–549 (2005)
26. B.M. Rosen, C.J. Wilson, D.A. Wilson, M. Peterca, M.R. Imam, V. Percec, Chem. Rev. **109**, 6275–6540 (2009)
27. A.P. Singh, A. Biswas, A. Shukla, P. Maiti, Sig Transduct Target Ther. **4**(33), 1–21 (2019)
28. F. Bahadori, A. Dag, H. Durmaz, N. Cakir, H. Onyuksel, U. Tunca, G. Topcu, G. Hizal, Polymers **6**, 214–242 (2014)
29. K.A. Kosakowska, B.K. Casey, S.L. Kurtz, L.B. Lawson, S.M. Grayson, Biomacromol **19**(8), 3163–3189 (2018)
30. Y. Muraji, T. Fujita, H. Itoh, D. Fujita, Microsc. Res. **1**, 24–28 (2013)
31. A. Kumar, Y.K. Tyagi, Materials Today: Proceedings. **43**, 250–255 (2021)
32. V. Brega, F. Scaletti, X. Zhang, Li-S. Wang, P. Li, Q. Xu, V. M. Rotello, S. W. Thomas, ACS Appl. Mater. Interfaces. **11**, 2814–2820 (2019)
33. H.-Y. Li, Y.-H. Chu, Molecules **25**, 104–115 (2020)
34. N. Dayyani, S. Khoee, A. Ramazani, Eur. J. Med. Chem. **98**(15), 190–202 (2015)
35. L. H. Urner, Y. B. Maier, R. Haag, K. J. Pagel, Am. Soc. Mass Spectrom. **30**:174Y180 (2018)
36. M. Wyszogrodzka, R. Haag, Biomacromol **10**, 1043–1054 (2009)
37. N. Dayyani, A. Ramazani, A. Khoee, SILICON **10**, 595–601 (2018)
38. C. Durukan, B. Kiskan, Y. Yagci, Polymers **11**, 679–690 (2019)
39. D. Lombardo, M. A. Kiselev, S. Magazù, P. Calandra, Advances in Condensed Matter Physics Volume **2015**, Article ID 151683 (2015)
40. Y. Wang, Y. Zhang, Xi. Liu, J. Wang, Li. Wei, Y. Feng, J Surfact Deterg. **17**, 295–301 (2014)
41. B. Lin, A.V. McCormick, H.T. Davis, R. Strey, J. Colloid Interface Sci. **291**(2), 543–549 (2005)
42. N. Bagalkot, A.A. Hamouda, O.M. Isdahl, MethodsX. **5**, 676–683 (2018)
43. Y. Gohon, J.D. Vindigni, A. Pallier, F. Wien, H. Celia, A. Giuliani, C. Tribet, T. Chardot, P. Briozzo, Biophysica Acta. **1808**, 706–716 (2011)
44. T. Yoshimura, A. Ebihara, H. Iwase, Colloids Surf. A **533**, 197–203 (2017)

Proceedings of the 6th International Conference on Communication and Electronics Systems (ICCES-2021)
IEEE Xplore Part Number: CFP21AWO-ART; ISBN: 978-0-7381-1405-7

Designing of Serial Asynchronous to Analog Resolver Format Converter

Vijendra Kumar Singh
M.Tech(VLSI), 4TH Sem
Department of Electronics and Comm.
Delhi Technological University, Delhi, India
Kvijendra.66@gmail.com

Dr. Yashna Sharma
Assistant Professor
Department of Electronics and Comm.
Delhi Technological University, Delhi, India
yashnasharma@dtu.ac.in

2021 6th International Conference on Communication and Electronics Systems (ICCES) | 978-1-6654-3587-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICCES51350.2021.9489231

***Abstract*: - This article presents the design of an interface solution for two systems where the input system gives angular information in serial asynchronous format while the end system requires it in Analog Resolver format. The angular information examined here was the ship's Roll and Pitch as obtained from the ship's RLG (Ring Laser Gyro) equipment. The OrCAD cadence proprietary software tool package was used to create an embedded board. The proposed model is developed by using microcontroller 8051, Digital to Resolver converter module and some other required components, and further it was tested by using simulated Roll Pitch angular information, which is generated with the help of COM Port Toolkit and the output of this interface solution was measured on API (Angle Position Indicator) and Oscilloscope.**

Keywords: - Line to Line Voltage, Serial asynchronous, Resolver, RS422, Roll, Pitch

1. Introduction

Old ships are fitted with various analog equipments/systems. Like, the mechanical gyroscope will be fitted in onboard ships to send Roll and Pitch angular information in Analog Resolver format to the end system. But as the part of modernization, Old mechanical Gyroscope was replaced by latest RLG (Ring Laser Gyroscope) [5]. So, the biggest challenge here in the integration of these two equipments is that, the latest RLG is not compatible with the old mechanical Gyroscope. The format in which the End System accepts ship's Roll and Pitch angular information is not available in same format in RLG. Therefore, an interface solution has been designed to integrate the latest RLG with End System [9].

2. Block Diagram of Proposed Interface Solution

The block diagram in Figure-1 shows the top level functionality of the interface solution and about modules used in the proposed design [2]. The driver IC AM26LS32 [14] is used to convert the RS422 input into TTL format [11]. The microcontroller 8051 [15] is used to extract the Roll and Pitch angular information from the Serial Asynchronous input packet [1] and fed the Roll and Pitch to DSC11524 Modules. These modules convert the angular information into Analog Resolver format [12]

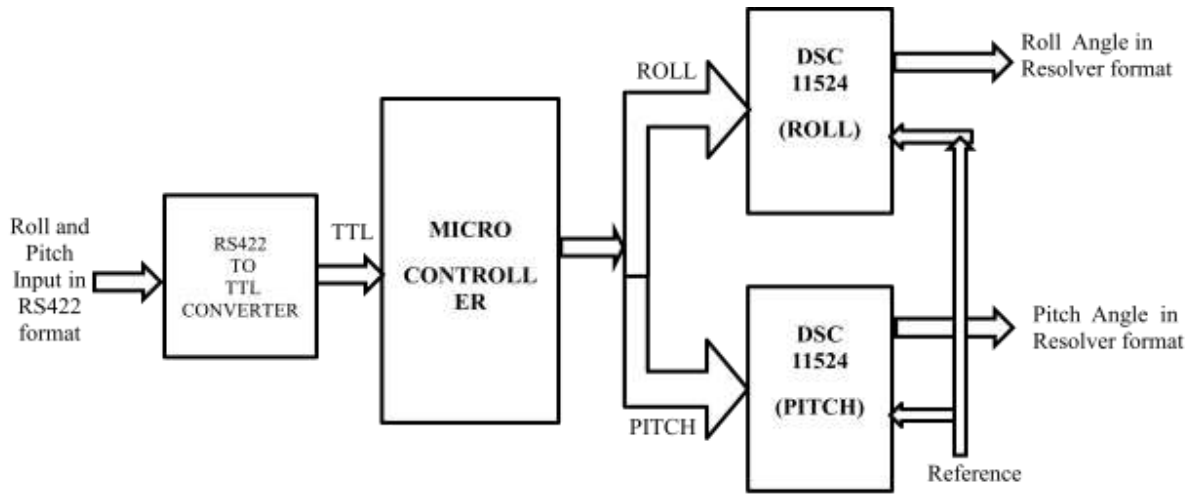


Fig. 1. Block Diagram of Proposed Interface Solution.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:08:46 UTC from IEEE Xplore. Restrictions apply.

3. The Interface Solution Design

The Roll and Pitch angular information of ship was converted [7] to the Analog Resolver format [3] by the designed interface solution. The key factors that led to this particular design [9] are as follows:-

(a) The design to have scalability to cater for change in requirements in future.

(b) The interface solution should be easy to maintain onboard.

(c) The interface solution should not add latency in the system.

(d) The software algorithm should be accurate and should not introduce errors in the angular information during conversion [12].

3.1.1 Circuit Diagram

The circuit diagram mentioned at Figure-2 was designed using OrCAD Capture cadence proprietary software tool suite [4]. OrCAD Capture is a powerful schematic drawing tool for implementing the design intent of an electronic circuit. The major components used in designing the circuit are the Microcontroller 8051 IC AT89C55, Digital to Resolver Converter IC DSC11524-303 (Two IC,s are used one for Roll Angle and another for Pitch angle), One IC for TTL to RS422 standard conversion AM26LS32, Four Low Power Operational amplifiers LM324 and some other driver IC,s.

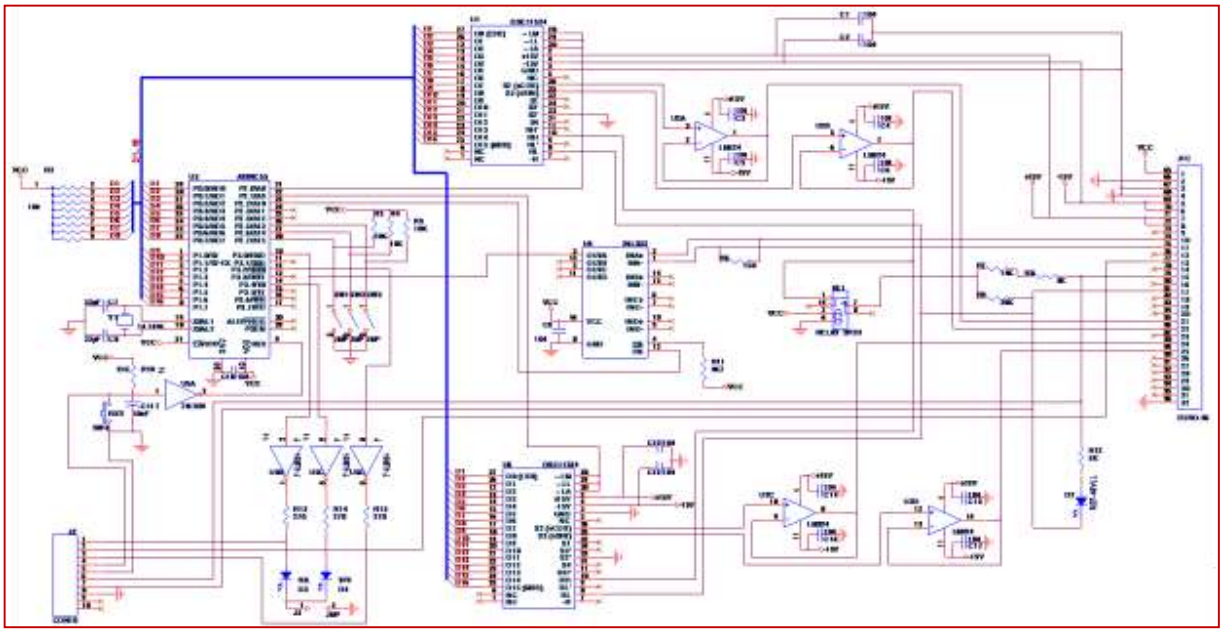


Fig. 2. Circuit Diagram of Proposed Interface Solution.

3.1.2 The Interface Solution Card

The PCB (Printed Circuit Board) of interface solution is mentioned at Figure-3. This PCB was designed by using the OrCAD PCB Editor Cadence software [10]. Two layer PCB is used for designing the required PCB. These are the basics steps used in PCB designing i.e. Draw the Outline of PCB, assign the Foot prints of all components, creating a netlist, automatic or manual Routing, selection of Top or Bottom layer, selection of constraints, generating the Artwork (Gerber file).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:08:46 UTC from IEEE Xplore. Restrictions apply.

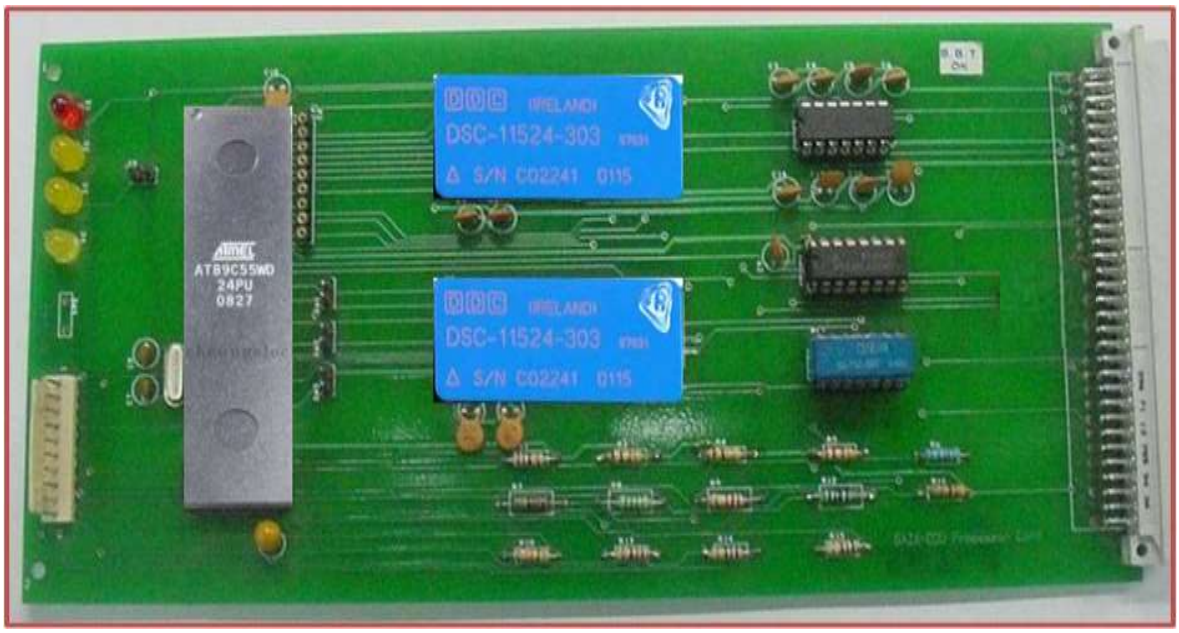


Fig. 3. Interface Card of Proposed Interface Solution.

4. Roll Pitch Angular Information Simulation for Testing

The interface solution has been tested with the help of simulated Roll Pitch angular information generated by COM Port Toolkit [6]. This COM Port Toolkit is PC based application that mimics the output behavior of the latest RLG of the Ship [10]. This application software provides a user friendly interface to configure RLG data. When we need to simulate RLG data, we write down the RLG packet structure in this tool, specify the appropriate serial parameters, and then click the transmit button. This serial asynchronous packet with previously defined serial parameters may then be received from any output USB port of the PC. A commercial USB to RS422 converter has been used in this simulation [6] because of the interface card, which needs the input data in RS422 format. This Rs422 format data fed to input of card and output of card was checked and then the performance of the interface card was evaluated. A snapshot of the COM Port Toolkit with simulated values is shown in Figure 4. This kind of validation philosophy helps to capture error/ discrepancies, if present, at an early stage of the design and development of interface solution, and thus saves a lot of time during the onboard integration stage [10].

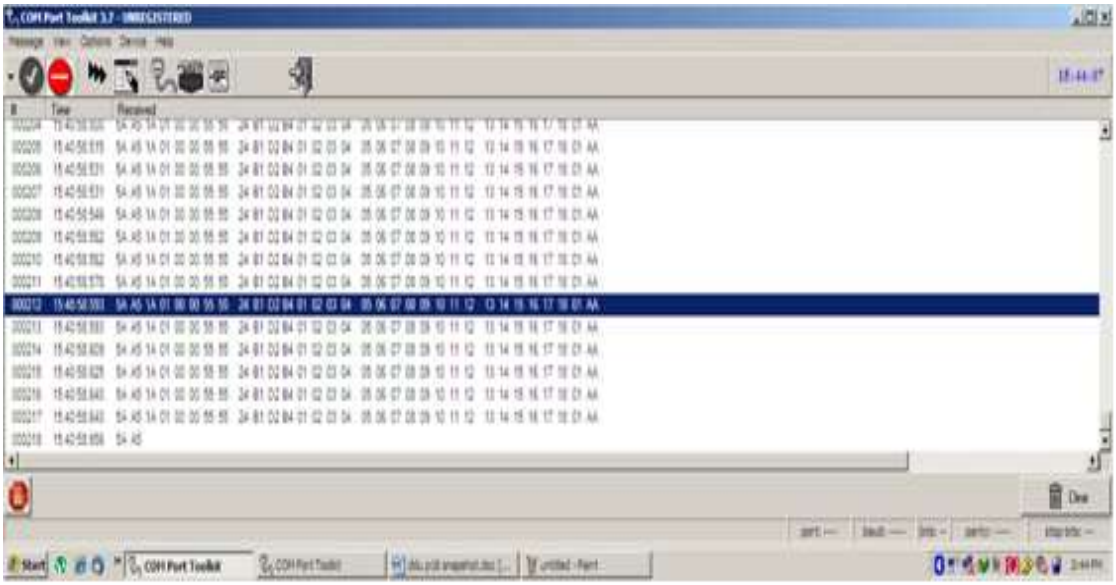


Fig.4 Screenshot of COM Port Toolkit with simulated Roll and Pitch Angular Information

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:08:46 UTC from IEEE Xplore. Restrictions apply.

5. Testing Setup and Testing

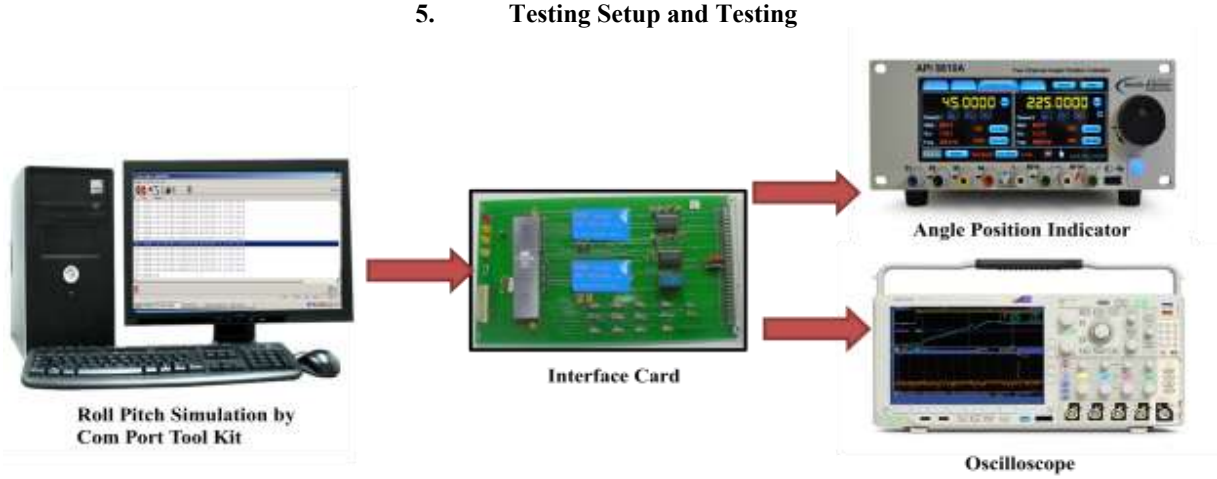


Fig. 5. Testing Setup for Interface Card.

The testing setup for interface solution is mentioned at Figure-5. The simulated values of Roll and Pitch angular information have been generated with the help of PC based COM Port Toolkit application [6] and fed to input of Interface card with the help of a USB to serial converter. The output of Interface card is measured on Angle Position Indicator and output waveforms have been captured on Oscilloscope.

6. Result

The Range of Roll and Pitch angle is ±90º, the clockwise angle measured as positive angle and in anticlockwise direction angle will be measured as negative. The angle measured at the output of interface solution will be measured by using 2X factor. It means the 90º angle displays on Dial in 180º scale. The 1º angle will be displayed in 2º scale, just double. So it is called as 2X factor. It is explained below:-

a) As per 2X factor, for positive angle like +45º will be represented on Dial as +45º*2=+90º, it means +45º will be represented as +90º on Dial.

b) As per 2X factor for negative angle like -45º will be represented on Dial as +360-45º*2=+270º, it means -45º will be represented as +270º on Dial.

The output of card is in Analog Resolver format. The voltage measured at the output of interface card will be measured line to Line; therefore it is called as VLL. It is RMS value. These voltage values for different Roll and Pitch angles were measured with the help of oscilloscope and mentioned in Table 1 and 2.

6.1.1 Output Voltage Calculation with the help of Oscilloscope

Output of interface card is in Resolver format. The Resolver signal travels on four lines/windings called as S1, S2, S3 and S4. S1 and S3 windings are called as sine windings and S2 and S4 are called as cosine windings. The voltage was measured between two lines S1, S3 and S2, S4 [13]. It is explained below:-

- Line to line Voltage measured between S1 and S3:-

  V(S1,S3)=11.8Volt*Sin(2*Roll or Pitch angle)

- Line to line Voltage measured between S1 and S3:-

  V(S2,S4)=11.8Volt*Cos(2*Roll or Pitch angle)

Output Resolver signal voltage was measured with the help of oscilloscope and same values were calculated with the help of above mentioned formulas [8] and values were matched.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:08:46 UTC from IEEE Xplore. Restrictions apply.

6.1.2 Roll Angle Measured on API and on Oscilloscope at the Output of Interface Solution

S. No.	Simulated Input Roll Angle	Output Roll Angle on API	Voltage(VLL)	
			S2,S4 Cos Winding	S1,S3 Sin Winding
1.	0	0	12.15	0.003
2.	1	1.94	12.14	0.218
3.	5	9.91	11.95	2.085
4.	7.5	14.07	11.72	3.107
5.	10	19.84	11.40	4.105
6.	15	29.83	10.51	6.005
7.	22.5	44.76	8.6	8.52
8.	30	59.83	6.044	10.44
9.	45	89.98	0.003	12.07
10.	60	120.16	6.040	10.43
11.	75	150.17	10.51	6.003
12.	90	180	12.15	0.002
13.	3.75	7.41	12.03	1.567
14.	1.875	3.69	12.11	0.788
15.	-1	358.02	12.13	0.416
16.	-7.5	345.11	11.72	3.105
17.	-11.25	337.63	11.21	4.593
18.	-15	330.13	10.51	6.004
19.	-22.5	315.18	8.59	8.54
20.	-30	300.13	6.039	10.44
21.	-45	270	0.003	12.06
22.	-60	239.80	6.04	10.44
23.	-75	209.81	10.52	6.004
24.	-90	180	12.14	0.003

Table 1. Roll angle measured at the output of Interface Solution.

6.1.3 Pitch Angle Measured on API and on Oscilloscope at the Output of Interface Solution

S. No.	Simulated Pitch Angle	Output Pitch Angle on API	Voltage(VLL)	
			S2,S4 Cos Winding	S1,S3 Sin Winding
1.	0	359.98	12.12	0.002
2.	1	1.94	12.12	0.415
3.	2.813	5.54	12.06	1.174
4.	5	9.89	11.94	2.079
5.	5.625	11.12	11.89	2.335
6.	10	19.84	11.39	4.098
7.	11.25	22.33	11.20	4.585
8.	15	29.81	10.50	5.995
9.	22.5	44.80	8.58	8.52
10.	30	59.82	6.035	10.42
11.	45	89.98	0.0025	12.04
12.	60	120.16	6.035	10.42
13.	75	150.15	10.50	5.999
14.	90	180.00	12.12	0.0045
15.	-1	358	12.12	0.419
16.	-1.1	357.81	12.12	0.460
17.	-10	340.11	11.40	4.103
18.	-15	330.13	10.50	6.501
19.	-22.5	315.15	8.59	8.53
20.	-30	300.16	6.037	10.43
21.	-45	269.98	0.005	12.05

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:08:46 UTC from IEEE Xplore. Restrictions apply.

22.		-60	239.80	6.037	10.42
23.	**T**	-75	209.78	10.50	5.997
24.	**T**	-90	179.96	12.13	0.003

a

bl Table 2. Pitch angle measured at the output of Interface Solution.

7. Conclusion

An interface solution has been designed, developed and tested with the help of simulated Roll and Pitch angular information in serial asynchronous packet, generated by COM Port Toolkit application software. The outputs have been measured on API (Analog Position Indicator) and output VLL (Line to Line Voltage) has been measured on oscilloscope. The Roll and Pitch angle values have been found correct with some minute deviation and the VLL voltage for different angle values have been found correct with some minute deviation. Here, the main module used for conversion of digital to Resolver has 16 bit resolution and ±4minutes accuracy and it can operate from 0 to 70ºC temperature. This Interface Solution could be utilized for interfacing of latest RLG and End System with some modifications.

References

1. Serial Port Complete, 2nd Ed. by Jan Axelson.
2. Hoang, N.; Jenkins, M. & Karangelen, N. Data integration for military systems engineering. In Proceedings of the IEEE Symposium and Workshop on Engineering of Computer Based Systems, 1996. pp. 11-15. doi:10.1109/ecbs.1996.494504
3. Khaleghi, B.; Khamis, A.; Karry, F. & Razavi, S. Multisensor data fusion: A review of the state-of-the-art. Information Fusion, 2013, 14(1), 28-44. doi: 10.1016/j.inffus.2011.08.001
4. Okamoto, S. & Sycara, K. Augmenting ad hoc networks for data aggregation and dissemination. In Proceedings of the 28th IEEE Conference on Military Communications, 2009, pp. 2346-2352. doi:10.1109/milcom.2009.5379977
5. Reddy, G.S. & Saraswat, V.K. Advanced navigation system for aircraft applications. Def. Sci. J., 2013, 63(2), 131-137. doi: 10.14429/dsj.63.4254
6. Popa, M., Popa, A. S., Cretu, V., & Micea, M. (2006). Monitoring Serial Communications in Microcontroller Based Embedded Systems. 2006 International Conference on Computer Engineering and Systems. doi:10.1109/icces.2006.320425
7. Lei, D., Chang-shi, X., Yuan-qiao, W., & Le, W. (2015). Safe speed of ships passing access channel of port. 2015 International Conference on Transportation Information and Safety (ICTIS). doi:10.1109/ictis.2015.7232076
8. Griffiths, G., Bradley, S. E., & Ruiz, S. (n.d.). Deep water bottom-track ship's velocities from an acoustic correlation current profiler. Oceans '97. MTS/IEEE Conference Proceedings. doi:10.1109/oceans.1997.624202
9. Del Amo, A., Montero, J., & Gomez, D. (2006). Fuzzy Logic Applications to Fire Control Systems. 2006 IEEE International Conference on Fuzzy Systems. doi:10.1109/fuzzy.2006.1681877
10. Defence Science Journal, Vol.64, No.5, September 2014, pp.445-450, DOI:10.14429/dsj.64.5035 "Designing Integration Unit to Integrate Navigational and Tactical Equipment onboard Naval Platforms".
11. Jie Liang and Ran Duan, "Design of the Embedded Serial and Parallel Communication Protocol Controller", IEEE, pp. 436-439, 2010.
12. Bhagiratharao, E. Real time information fusion in military systems. Def. Sci. J., 1990, 40(1), 71-82.
13. Yongjian Tang, Lenian He and Xiaolang Yan, "A novel data processing in high speed serial communication", IEEE, pp. 1228-1231, 2005.
14. Texas Instruments datasheet "onAM26LS32AC, AM26LS32AI, AM26LS32AM AM26LS33AC, AM26LS33AM" SLLS115F – OCTOBER 1980 – REVISED SEPTEMBER 2016
15. Book "8051 Microcontroller and Embedded Systems using assembly and c", second edition, Muhammad Ali Mazidi.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:08:46 UTC from IEEE Xplore. Restrictions apply.

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.

Drying Kinetics and Economic Analysis of Bitter Gourd Flakes Drying Inside Hybrid Greenhouse Dryer

Asim Ahmada
Birla Institute of Technology
Om Prakash (16omprakash@gmail.com)
Birla Institute of Technology
Anil Kumar
Delhi Technological University

Research Article

Keywords: Bitter gourd flakes, Hybrid greenhouse dryer, CO2 mitigation, Drying kinetics, Thermal Storage.

DOI: https://doi.org/10.21203/rs.3.rs-683304/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Drying Kinetics and Economic Analysis of Bitter Gourd Flakes Drying inside Hybrid Greenhouse Dryer

Asim Ahmad[a,b], Om Prakash[a*], Anil Kumar[c*]

[a]Department of Mechanical Engineering, Birla Institute of Technology, Mesra, Ranchi, India

[b]Faculty of Engineering & Applied Sciences, Usha Martin University, Ranchi, India

[c]Department of Mechanical Engineering, Delhi Technological University, Delhi, India

[a*]Corresponding Author E-mail ID: 16omprakash@gmail.com

[c*]Corresponding Author E-mail ID: anilkumar76@dtu.ac.in

ABSTRACT

In this study, heat storage-based hybrid greenhouse dryer has been developed and analysed for drying bitter gourd flakes under the climatic condition of Ranchi, India. The propose heat storage based hybrid greenhouse dryer consists of a solar air heater with 2.12 m^2 area, greenhouse dryer and DC fan to induce and force the air. The significant objective of the present study is to analyse the drying efficiency, drying kinetics, property analysis, economic analysis, embodied energy and CO_2 mitigation of the hybrid greenhouse dryer for drying of bitter gourd flakes. An experiment was performed simultaneously on proposed system and open sun drying for the proper comparative analysis. The results show that the mass of crop reduced from 4000g to 458g in 6 hr in proposed dryer and 4000g to 500g in 15 hr for open system. Thus, significant drying time is reduced in proposed system by 8 hours as compared to open system. Environmental impact analysis shows that the EPBT was found to be 0.4907 years only. The cost analysis of the proposed system dryer was 22664.30 INR. The total embodied energy is found to be 1591.07 kWh and earned carbon credit ranges from 16844.76 INR to 67379.05 INR, while CO_2 mitigation was 46.28 tonnes for 35 years of expected lifetime. Seven standard mathematical models for drying of bitter gourd flakes were studied. Ahmad and Prakash model was found to be best as compared to other models. The metal contents of dried bitter gourd flakes were also examined. Bitter gourd dried in proposed dryers possesses superior metal content as compared to open system. Impact analysis demonstrates that the hybrid greenhouse dryer is more suitable for reducing post-harvest loss with environmental sustainability.

Keywords: Bitter gourd flakes, Hybrid greenhouse dryer, CO_2 mitigation, Drying kinetics, Thermal Storage.

INTRODUCTION

Bitter gourd is an essential tropical crop that comes from the Cucurbitaceous family. It is mainly found in countries like China, India, Thailand, Malaysia and Japan during the rainy season. Bitter gourd is famous for its high nutrients value and a good source of iron and phosphorous. Bitter gourd consists of high moisture content, i.e. 88%-93%, leading to rapid microbial spoilage and biochemical changes (Chauhan et al., 2018). Therefore, various food preservation methods are used to preserve it like conventional drying, canning and freezing. However, all these techniques are energy-intensive and costly for small farmers. Open sun drying is a traditional method for drying bitter gourd for a long time, especially in rural areas. Still, unexpected rain and foreign particles contamination make it unsafe and unhygienic (Abubakar et al., 2018 a). These issues can be resolved by applying a controlled mode of solar drying. Solar dryer is an example of the controlled mode of solar drying. The significant difficulty faces by the solar dryer is low drying efficiency and dependency on ambient parameters. Hence, the thermal storage concept is crucial for minimizing the fluctuation of ambient parameters and increasing the drying efficiency by enhancing drying time (Abubakar et al., 2019).

Greenhouse dryer is a type of direct solar dryer. It helps in drying the crop in a larger quantity. Greenhouse dryer is an appropriate dryer for low and medium temperature thermal drying. Many researches have been carried out on the development of greenhouse dryer and performance enhancement. To stop heat loss from the north wall, many researchers proposed the idea to which losses can be minimised from the north wall greenhouse (Abubakar et al., 2018 b). The main disadvantage in greenhouse dryer is that crops cannot be dried in off-sunshine hour (Prakash et al., 2016). To minimize this issue, the greenhouse dryer can be integrated with heat storage systems (Čiplienė et al., 2015; Kant et al., 2016). In order to improve the efficiency of greenhouse dryer, it can be coupled with solar air heater to make as hybrid system. The additional heat storage materials help to provide constant heat throughout the drying process particularly in off sunshine hour. A hybrid solar dryer system integrated with thermal energy storage (rock bed) was tested for the drying of Guduchi, ginger and turmeric (Prasad and Vijay 2005). The drying analysis was done on open sun drying, solar dryer and solar dryer with biomass. Drying efficiency was found to be highest by the author in a hybrid dryer in comparison to open sun drying and solar dryer. The indirect solar dryer operating in active mode integrated with TES (wax-based) was evolved by (Shalaby and Bek 2014). The experimental analysis was done for drying of two herbs, i.e. T. Neriifolia and O. Basilicum. In this developed system, the heat was supplied to drying chamber at the off sunshine hour with the help of thermal storage material. Hence final moisture content was achieved in 18 hours

and 12 hours for drying of T. Neriifolia and O. Basilicum. An indirect solar dryer operating under passive mode integrated with multiple solar air heaters was developed for drying of mint. The drying efficiency was reported as 28.2% based on total evaporation of water content and energy received on the area of the collector (Jain and Tewari 2015). The experiment was performed in passive greenhouse dryer in different floor conditions by Ahmad and Prakash. The different floor condition is developed based on the different types of sensible thermal storage material (Gravel, Concrete and Black painted gravel). The drying efficiency of the passive greenhouse dryer with black painted gravel is highest for drying of tomato flakes as compared to other sensible thermal energy storage material (Ahmad and Prakash 2020; Ahmad and Prakash 2021).

From the intense literature survey, it is found that very little research happens in the field of solar dryer coupled with thermal storage concept to dry the agricultural produce (Khanlari et al., 2020). However, there is no research published on drying bitter gourd flakes using mixed-type thermal storage-based hybrid greenhouse dryers.

Hence, the drying experimentation (bitter gourd flakes) is being conducted in two different modes: hybrid greenhouse dryer under mixed thermal storage and open sun drying. The drying performance (drying efficiency, moisture ratio, moisture diffusivity, moisture content), drying kinetics, energy analysis (embodied Energy, carbon credit, carbon emission, energy payback period), and payback period by the cost of the proposed dryer is being evaluated. Further, the quality and valuable metallic properties of the bitter gourd flakes dried in the proposed dryer and open sun drying are compared with fresh samples.

EXPERIMENTAL SETUP

In the proposed hybrid greenhouse dryer, Greenhouse dryer is coupled with a single pass solar air heater (SAH). The wall of the dryer is covered with a polycarbonate sheet of thickness 3 mm. The dimensions of the greenhouse dryer were taken as per the standard dimension, i.e. 1.5 m×1 m×0.5 m. In order to achieve maximum solar radiation, the roof of the dryer and SAH is tilted at a latitude angle of Ranchi, India, i.e. 23.34^{o}N. The two 12V DC fan of variable speed powered by 40W solar panel is placed to ventilate the humid air out of the greenhouse dryer and to induce preheated air from SAH to pass inside the dryer. The north wall of the greenhouse dryer is insulated with the combination of reflective mirror and thermocol. The mirror reflects the outgoing radiation from the north wall to inside drying space of the dryer. The thermocol reduces the heat transfer loss from the mirror to the surrounding. The mixed type thermal heat storage material, i.e. combination

of latent and sensible heat storage is applied at the floor of the dryer. The paraffin wax of 35 kg is used as latent heat storage material, while black-painted gravel of 29 kg is used as a sensible heat storage material. The paraffin wax is kept within the leak proof aluminium jacket at the ground of the greenhouse dryer covered with black painted gravel to provide proper thermal storage. The detailed thermo-specification of both thermal storage materials (TSM) is shown in Table 1. Initially Black painted gravel gets charge then it leads to charge the paraffin wax. The thermal storage material (TSM) gets charge from sunrise to noon when solar intensity increases continuously, as the solar intensity decreases discharging process gets starts. The discharging process will continue up to few post sunset hour. The heat release from TSM helps to maintain constant drying process during low solar intensity.

The solar air heater (SAH) is used provide the preheated air inside the drying chamber. The SAH is made up of wooden frame and have the dimension of 2.04 m×1.04 m×0.5 m. For the coupling of SAH with greenhouse dryer, a well insulated duct of diameter 0.5 m made of galvanized iron is being used. The insulating material (polyurethane foam) is used to minimize the heat loss from the connecting duct to surrounding. A DC fan is placed at the outlet of the SAH to provide pre heated air from the SAH to the drying chamber through duct. A black painted galvanized iron sheet is used as an absorber plate in a SAH. The SAH is covered with transparent glass of thickness 3mm. The structure of solar air heater is supported by L shaped Mild steel angle. The exposed part of the structure is painted to protect it from adverse environmental effects. The detailed schematic diagram with proper dimension and pictorial view of the experimental setup are explained in Fig 1(a & b).

MATERIALS & METHODS

SAMPLE PREPARATION

The author aimed to evaluate the performance of the proposed system in load condition for drying of bitter gourd flakes. The experiment was conducted for two consecutive days, i.e. 11-13 May 2021, from 10 AM morning to 5 PM evening in BIT Mesra, Ranchi, Jharkhand, India for drying of crops in hybrid mode and open mode. The fresh bitter gourde of 4 kg were purchased from the local market of Ranchi were washed and dried with cotton clothes. This bitter gourd were sliced into 1–2 mm thick flakes and taken for the experimentation.

INSTRUMENT

Suitable instruments were used to observe the crucial parameters on an hourly basis like ambient temperature, relative humidity, global radiation, wind velocity, mass of crops, crop temperature, outlet temperature and room temperature. The details of the instruments are shown in Table 2.

UNCERTAINTY ANALYSIS

Experimental errors come from the environmental conditions, selection of the instruments and the measured value. This analysis is very appropriate to estimate the error between the estimated and observed parameters. The uncertainty analysis for the dependent parameters can be evaluated by Eq. 1 (Ahmad and Prakash 2020).

$$\Delta p = \sqrt{\left(\frac{\partial p}{dk_1}\Delta k_1\right)^2 + \left(\frac{\partial p}{dk_2}\Delta k_2\right)^2 + \ldots\ldots\ldots\ldots + \left(\frac{\partial p}{dk_n}\Delta k_n\right)^2} \quad (1)$$

Where, p is the dependent parameters and k_1, k_2,…k_n is the measured value.

The total uncertainty of the experiment was evaluated as ±0.68%, which is under the permissible range. The detailed Uncertainty of the experiment is shown in Table 3

NUMERICAL ANALYSIS

Moisture Ratio (M_{ratio})

The moisture ratio can be defined as the ratio of the hourly moisture content to the initial moisture content of the crop. The initial moisture content (M_{ini}) for bitter gourd flakes on the wet basis (w.b.) can be calculated as

$$M_{ini} = \frac{W_{ini} - W_{hr}}{W_{ini}} \times 100 \quad (2)$$

Where, W_{ini} is the initial weight of bitter gourd flakes, W_{hr} is the hourly weight of bitter gourd flakes. The final and initial moisture content help to determine the total removal of water content (W_{ttl}) from bitter gourd flakes (Fudholi et al., 2014)

$$W_{ttl} = \frac{M_{ini} - M_{fnl}}{100 - M_{fnl}} \times W_{ini} \quad (3)$$

Where M_{ini} the initial moisture is content and M_{fnl} is the final moisture content on w.b.

The instantaneous moisture content on d.b. can be calculated as (El-Sebaii and Shalaby 2013)

$$M_{ins} = \left[\frac{(M_{ini} + 1) W_{hr}}{W_{ini}} - 1 \right] \quad (4)$$

The moisture ratio (M_{ratio}) can be calculated as (Akpinar, E.K., 2010; Sethi and Dhiman 2020)

$$M_{ratio} = \frac{(M_{ins} - M_{eqm})}{(M_{ini} - M_{eqm})} \quad (5)$$

Where M_{eqm} is the equilibrium moisture content and M_{ini} is very high as compared to M_{eqm}. Thus, equilibrium moisture content can be neglected. Hence, the equation can be expressed as

$$M_{ratio} = \frac{M_{ins}}{M_{ini}} \quad (6)$$

Drying efficiency (η_{hgd})

Drying efficiency is the ratio of the heat utilize for drying of crop to the total energy consumed by the dryer. Drying efficiency depends on the parameters like latent heat, total removal of water, area of tray and global solar radiation. It can be evaluated as Eq. 7. (Banout et al., 2011; Prakash and Kumar 2014; Doymaz, I. 2007):

$$\eta_{hgd} = \frac{W_{ttl} \times h_{lh}}{A_{tr} \times I_{grd}} \quad (7)$$

Embodied Energy & Cost Analysis

The total amount of energy used for the production of any product or service is termed as embodied energy. The embodied energy for the various component used in the fabrication of a hybrid greenhouse dryer is represented in Table 4 (Banout et al., 2011; Prakash and Kumar 2014; Doymaz, I. 2007).

The total cost incurred for the fabrication of the proposed system is INR 22664.30; it majorly consists of polycarbonate sheet, paraffin wax and polycrystalline solar cell. The details of cost of the various component used in the proposed system is shown in Table 5.

Energy Payback Time (EPBT)

The time requires to recover the embodied energy of the system or service is called energy payback time. The dryer payback period is evaluated as (Banout et al., 2011)

$$EPBT = \frac{E_m}{AE_o} \tag{8}$$

The annual energy output (AE_o) is the product of daily energy output (DE_o) and the number of working days (N_{day}).

$$AE_o = DE_o \times N_{day} \tag{9}$$

$$DE_o = \frac{W_{ttl} \times h_{lh}}{3.6 \times 10^6} \tag{10}$$

$$DE_i = I_{grd} \times N_{hr} \times A_{rm} \times 10^{-3} \tag{11}$$

Where E_m is the Embodied Energy of Hybrid greenhouse dryer, AE_o is annual energy output, DE_o is daily energy output, DE_i is daily energy input. According to the climatic condition in Ranchi, India, there are 300 days as an annual average sunny day, which is represented by N_{day} and 1800 hours as annual average sunny hours and seven sunny hr on a daily basis which is represented as N_{hr}.

CO_2 Emission

The average emission of CO_2 is equivalent to 0.98 kg of CO_2 per kWh for the coal based electricity generation. CO_2 emission per year ($CO_{2\ em/yr}$) depends on the parameters like embodied energy and lifetime of the system (LT). CO_2 emission rate/yr can be evaluated as (Prakash and Kumar 2014)

171 $$CO_{2\ em/yr} = \frac{E_m \times 0.98}{LT} \quad (12)$$

172 ## Effective moisture diffusivity

173 The amount of movement of moisture in the crops during drying process is known as effective moisture
174 diffusivity. Effective moisture diffusivity is affected by the porosity of the crop, temperature, moisture content
175 and compositions. It can be evaluated as per Eq. 13 or Eq. 14.

176 $$MR = \frac{M_t - M_e}{M_o - M_e} = \frac{8}{\pi^2}\exp\left(-\frac{\pi^2 D_{eff}}{4r^2}gt\right) \quad (13)$$

177 Where D_{eff} is the actual diffusivity coefficient (m^2/s), t is the drying time, r is the thickness of the slab (m), and n
178 is the positive integer. Only the first term of Equation (13) can be used for a long drying period [25]: Many
179 researchers demonstrated that Equation (13) could be further simplified to a straight-line equation as Eq. 14
180 (Lopez et al., 2000; Sacilik et al., 2006):

181 $$\ln(MR) = \ln\left(\frac{8}{\pi^2}\right) - \left(\frac{\pi^2 D_{eff}}{4r^2}gt\right) \quad (14)$$

182 The actual moisture diffusivity (D_{eff}) is typically determined by plotting experimental drying data in terms of ln
183 (MR) versus time and the slope (k_o) as shown in Eq. 15 (Wang et al., 2004):

184 $$k_o = \frac{\pi^2 D_{eff}}{4r^2} \quad (15)$$

185 where r is the thickness of drying crop in mm, i.e. bitter gourd flakes in our case.

186 ## Drying Kinetics

187 The drying kinetics have been analysed to dry bitter gourd flakes inside hybrid greenhouse dryers and open
188 mode drying. The first model is developed by Newton for thin layer types of solar drying. The variable k is the
189 exponential variable in this model, which is taken as constant. Another exponential equation is a page model.
190 Further, the up-gradation of the Newton model is developed by Henderson and Pabis consist of coefficient a.
191 The up-gradation of Henderson and Pabis is the logarithmic model, which consists of an extra constant c.

Further Wang and Singh model is proposed, which is a quadratic function, and the performance of this function is better than the previous models. A suitable selection of the drying model is necessary to describe the drying process. Table 6 represents the seven drying models for drying bitter gourd flakes. The results of the investigation were fitted in the mathematical model mentioned in Table 6 with the help of Matlab 13a using non-linear regression. For the validation of the statistical analysis, the drying kinetics has been settled on the basis of Sum of square (SSE) due to error, degree of freedom Adj R^2, Root Mean Square error(RMSE) and R square.

Metallic concentration of dried bitter gourd flake:

Dried bitter gourd flakes under both modes of drying were converted into diluted solution by microbial digestion. After dilution it was mixed with Nitric acid (HNO_3) and tested in the Central instrumentation facility (CIF) BIT Mesra. Six important metal concentration has been measured namely Sodium (Na), Potassium (K), Manganese (Mn), Magnesium (Mg), Iron (Fe) and Copper (Cu).

Result and Discussion

Variation of Ambient Parameters during experimentation.

Ambient parameters like solar radiation, air temperature, relative humidity and wind speed were taken on an hourly basis during experimentation. These parameters are presented in Fig. 3-5. Ambient parameters on both days were almost similar. Variation of solar radiation was observed as 660-1050 W/m^2 during the first day and 900-1072 W/m^2 during the second day while the average value of solar radiation was found as 936 W/m^2 and 932 W/m^2 on the first and second day respectively. Relative humidity is inversely proportional to ambient temperature; as ambient temperature increases, relative humidity decreases. The maximum ambient temperature was found at 13th hour in all the days of the experimentation. For open and hybrid mode, the maximum ambient temperature was found to be 36.35 ^oC and 40.90 ^oC. For open and hybrid mode the minimum relative humidity on day 1 was 38.20% and 29.90% while on day 2 it was 45.20% and 40.70% respectively. The variation in wind speed during the both day of experimentation was remains constant, i.e. 0.29-0.45 m/s.

Effect of SAH outlet temperature, Ground and Room Temperature during drying of bitter gourd flakes

Room and ground temperature of the proposed dryer is shown in Fig.6-7. The higher solar radiation leads to a high rise in-ground and room temperature of the proposed dryer. During sunshine hours, heat is absorbed by thermal storage materials applied on ground of the dryer. In off sunshine hours, stored heat gets released inside the drying chamber, which helps in maintaining the favourable drying temperature. The maximum ground temperature for proposed dryer on day 1 was 77.20^{o}C while on day 2 it was 71.40 ^{o}C. The maximum room temperature for proposed dryer on day 1 was 67.20 ^{o}C while on day 2 it was 61.40 ^{o}C. The SAH outlet air temperature is an important factor for providing pre heated air inside the drying chamber of present designed system. The maximum outlet temperature of SAH is reported as 37.6^{o}C which is 2.7^{o}C more than ambient/inlet temperature.

Effects on Moisture Ratio, Moisture Contents and Effective moisture diffusivity for drying of bitter gourd flakes.

In this section, the effect on moisture content, moisture ratio and actual moisture diffusivity during bitter gourd flakes drying were discussed. Fig. 8 represents the variation of moisture ratio and moisture content with respect to drying time. The final moisture content of bitter gourd flakes (0.78%) reached in hybrid system in 6 h at 0.063391 kg/s air mass flow rate. For open system final moisture content of bitter gourd flakes (9.10%) achieved in 14h. The drying time in the proposed dryer is low as compared to open sun drying due to supply of pre heated air from SAH. During the initial phase of drying, higher removal of moisture takes place because of high moisture content. As the drying time increases, the moisture evaporation rate gets decreases with the decrease in the moisture content of bitter gourd flakes.

The effective moisture diffusivity for bitter gourd flake drying under open and hybrid mode was evaluated by Eq. 12. The effective moisture diffusivity at two different mass flow rates can be presented by the graph between drying time and ln (MR). Fig. 9 shows the variation of ln (MR) and drying time in different modes, i.e. hybrid mode and open mode. The effective moisture diffusivity of bitter gourd flakes was found to be 11.650 x 10^{10} m^2/s for hybrid mode and 0.8666 x 10^{10} m^2/s for open mode. It yields the maximum value at the mass flow rate of 0.06147 kg/s due to the fact that the drying time is slower than other flow rates.

Effects on Embodied Analysis, Cost analysis and CO_2 Emission of Hybrid Greenhouse Dryer

Table 4 represents the embodied energy analysis of the proposed dryer for the drying of bitter gourd flakes. The fabrication of the hybrid system is done with the help of locally skilled workers by using locally available materials. The cost of heat storage based hybrid greenhouse dryer was 22664.30 INR; it majorly consists of polycarbonate sheet, paraffin wax and polycrystalline solar cell. The detailed cost of the material used to fabricate hybrid greenhouse dryer is shown in Table 5. Various materials were used for the fabrication of the proposed system is also shown in Fig.10. The emission rate of CO_2/annum for the hybrid greenhouse dryer was reported as 33.86 kg, which is quite low as compared to the existing hybrid solar dryer (Prakash and Kumar 2014).

EPBT of Hybrid Greenhouse Dryer & Quality of dried crops

The energy payback period of the greenhouse dryer was calculated as per Eqn.9. The energy payback period of the developed hybrid greenhouse dryer is 1.87 year. The lifetime of the hybrid greenhouse dryer was found as 35 years; hence, the payback period is found very low as compared to the lifetime of the system.

To analyse the quality of dried bitter gourde, the dried bitter gourde sample of 0.5g was taken and placed in the burning cup. This cup is added with 15 ml of HNO_3 (pure form). Further, it is incinerated in the microwave oven at 200°C, and it is diluted by distilled water (50 ml). ICP-OES analysed mineral concentration with 3 seconds (copy time), 1-5 second (reading time), 5 to 12 mm (viewing height), 10.5-15 L/min (rate of flow of Plasma gas) and 0.8-1.4 kW (R.F Power). Mn and Mg types of minerals are highest in sun-drying due to the formation of ferric oxide from a chemical reaction that takes place in the microwave oven. The material concentration is shown in Table 7. The metal concentration of Na and K comes under the range of Suarez at al. while other metal like Mn, Mg, Fe and Cu is higher in our case. This is because various factors influence metal concentration like production region, cultivation method, cultivator and sampling period (Arslan and Özcan 2011). These metal contents also depend on the types of crops used for drying, soil quality of the crops, the water quality of the crops, Hence it varies as per the geographical locations (Suárez et al., 2007).

Drying Kinetics

The regression analysis has proceeded under hybrid mode and open mode for 7 thin-layer drying model for correlation of drying time and moisture ratio at the mass flow rate of air 0.06147 kg/s as shown in Table 6. Ahmad & Prakash model is the extension of the Logarithmic model. The equation describing Ahmad & Prakash model has four constant a, b, k and n. The variable t describes the time required for drying. Ahmad & Prakash model was selected as one of the best technique of curve fitting for analysis of non-linear regression based on R-square, adjusted R- square, SSE and RMSE. The statistical parameter values of the Ahmad & Prakash model were SSE 0.0186, R-square 0.9861, adjusted R-square 0.9823, and RMSE 0.04112. For the open sun mode, Ahmad & Prakash model was again found to be the best fitting curve model for non-linear regression analysis with SSE 0.01804, R-square 0.987, adjusted R-square 0.9835, and RMSE 0.0405. The Table 8 and 9 summarises all the thin layer models statistical analysis results for hybrid greenhouse and open sun dryer condition. The moisture ratios for the experimental and predicted values comparison are done for the confirmation of model selection for greenhouse dryer shown in Fig. 11 and for the open sun in Fig. 12. The fittingness of the statistical models is specified by the predicted values connected by a dotted line.

Efficiency of Hybrid Greenhouse dryer for drying

The daily thermal efficiency of a hybrid greenhouse dryer varies with the dryer's annual thermal Output and input energy. The daily thermal efficiency is found to be 50.20 %, while the annual thermal output is calculated as 732.20 kW, and input energy is found as 4.80 kWh. The experimental result shows that, as the solar radiation increases, the drying rate also increases, resulting in an increase in thermal efficiency. The thermal storage at the ground of the dryer provides steady heat inside the drying chamber, which results in large amount of transfer of heat to the crop in off sunshine hour. The mass flow rate of air is also an important parameter that affects the efficiency of the drying system. The increase in mass flow rate of air, results in the decrease of outlet temperature due to less contact duration of air inside the drying chamber. Hence it is necessary to maintain proper mass flow rate of air for enhanced efficiency.

Comparative Analysis of Hybrid Greenhouse Dryer with other Invigilator

The drying of bitter gourd flakes in hybrid greenhouse dryer compared with forced convection indirect greenhouse dryer (Vijayan et al., 2020). It was observed that a hybrid greenhouse dryer is more efficient in comparison to a forced convection indirect greenhouse dryer. This is because the drying time required for getting a stagnant crop weight is 6 h i.e. final moisture content as 7.8% for the proposed hybrid greenhouse dryer, while for the forced convection indirect greenhouse dryer used by it reached up to 9% in 7 h. The hybrid greenhouse dryer is more efficient because preheated air comes into the dryer section from SAH; also, the hybrid thermal energy storage concept is applied on the ground of the proposed dryer, which gives a proper distribution of heat inside the dryer. The detail comparative analysis of different parameters is shown in Table 10.

CONCLUSION

The present study considers the techno-environmental analysis & drying kinetics of hybrid heat storage based proposed dryer. The paraffin wax is being used as latent storage material, while black-painted gravel is used as sensible storage material on the ground of the drying chamber in the proposed system. The proposed hybrid system is the combination of the active mode greenhouse dryer and single-pass solar air heater. The bitter gourd flakes were dried in a Hybrid greenhouse dryer and simultaneously in natural (open sun) drying. The results observed by the above experimentation are as follows:

- The moisture content of bitter gourd flakes was decreased from 88.64% to 0.78% in the hybrid system, while in the open system, it reduces from 88.64% to 54.19% in 6 hours.
- Thus, the reduction in moisture content is 98.56% more in hybrid greenhouse dryer in comparison to open mode.
- The daily drying efficiency of the proposed hybrid system is 50.20 % at 0.06147 kg/s of the mass flow rate of exhaust air.
- The effective moisture diffusivity of bitter gourd flakes was found to be 11.650×10^{-10} m^2/s for hybrid mode, whereas, for open mode, it was 0.8666×10^{-10} m^2/s.
- The embodied energy and CO_2 emission per year of the proposed system were 1591.07 kWh and 33.86 kg.

- The cost analysis of heat storage based hybrid greenhouse dryer was 22664.30 INR; it majorly consists of polycarbonate sheet, paraffin wax and polycrystalline solar cell.
- After drying, both the modes' dried product was tested in the laboratory to find out the important metal content in it. The product dried in the proposed system was found to be higher metal content as compared to the open sun-dried product.
- The energy payback time for the proposed system was 1.88 years.
- The Ahmad and Prakash model and logarithmic models were selected as the best drying model for drying bitter gourd flakes in a Hybrid greenhouse dryer and open sun drying, respectively.

The heat storage based Hybrid greenhouse dryer was more consistent in drying and produces better quality dried products than open sun drying. Since the hybrid greenhouse dryer consumes less drying time than open sun drying, i.e. 8 hr for drying bitter gourd flakes, the heat storage-based hybrid greenhouse dryer is economically viable for small scale farmers and industries.

Acknowledgement:

The Author Acknowledges CIF, BIT, Mesra, Ranchi for providing support in this research.

Author contribution:

Asim Ahmad: Conceptualization, Methodology, Investigation, Data curation, Writing - original draft.

Om Prakash: Validation, Supervision, Writing - review & editing.

Anil Kumar: Investigation.

Funding: Not Applicable

Availability of data and materials:

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declaration of competing interest:

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Competing interests: The authors declare no competing interests.

Nomenclature

A_{rm} = Area of room (m^2)	$EPBT$ = Energy Payback time (Yr)	M_{ins} = Instantaneous Moisture Content
AE_o = Annual Energy Output (W)	E_m = Embodied Energy (kWh)	h_{lh} = Latent heat of evaporation (KJ/kg)
A_{tr} = Area of tray (m^2)	$CO_{2\ emmision/yr}$ = Emission rate of CO_2 per year (kg)	LT = Life time (Yr)
N_{day} = Average annual sunny days	M_{eqm} = Equilibrium Moisture Content	M_{ratio} = Moisture ratio
η_{hgd} = Drying system efficiency (%)	M_{fnl} = Final Moisture content	I_{grd} = Solar Radiation (W/m^2)
DE_o = Daily Energy output (W)	W_{hr} = Hourly Weight of Bitter gourd (kg/h)	W_{ttl} = Total removal of water content (kg)
DE_i = Daily Energy input (W)	M_{ini} = Initial Moisture Content	
N_{hr} = Daily sunny hour (hr)	W_{ini} = Initial Weight of Bitter gourd (kg)	

Abbreviation

eqm = equilibrium	hr = hourly	pbt = payback time
fnl = final	ini = initial	rm = room
grd = global radiation	ins = instantaneous	tr = tray
hgd = hybrid greenhouse dryer	lh = latent heat	ttl = total

355 **REFERENCES**

356 Abubakar, S., Umaru, S., Anafi, F.O., Abubakar, A.S. and Kulla, D.M., (2018 a). Design and performance
357 evaluation of a mixed-mode Solar Crop Dryer. FUOYE Journal of Engineering and Technology, 3(1),
358 pp.22-26.

359 Abubakar, S., Anafi, F.O., Kaisan, M.U., Narayan, S., Umar, S. and Umar, U.A., (2019). Comparative analyses
360 of experimental and simulated performance of a mixed-mode solar dryer. Proceedings of the Institution
361 of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 234(7), pp.1393-1402,
362 https://doi.org/10.1177%2F0954406219893394.

363 Abubakar, S., Umaru, S., Kaisan, M.U., Umar, U.A., Ashok, B. and Nanthagopal, K., (2018 b). Development
364 and performance comparison of mixed-mode solar crop dryers with and without thermal storage.
365 Renewable Energy, 128, pp.285-298, https://doi.org/10.1016/j.renene.2018.05.049.

366 Ahmad, A. and Prakash, O., (2019). Thermal analysis of north wall insulated greenhouse dryer at different bed
367 conditions operating under natural convection mode. Environmental Progress & Sustainable Energy,
368 38(6), p.p.1-12, https://doi.org/10.1002/ep.13257.

369 Ahmad, A. and Prakash, O., (2020). Performance Evaluation of a Solar Greenhouse Dryer at Different Bed
370 Conditions Under Passive Mode. Journal of Solar Energy Engineering, 142(1),pp.1-10,
371 https://doi.org/10.1115/1.4044194.

372 Ahmad, A. and Prakash, O., (2021). Development of mathematical model for drying of crops under passive
373 greenhouse solar dryer, Materials Today: Proceedings, ISSN 2214-7853, pp.1-4,
374 https://doi.org/10.1016/j.matpr.2021.05.180.

375 Akpinar, E.K., (2010). Drying of mint leaves in a solar dryer and under open sun: modelling, performance
376 analyses. Energy conversion and management, 51(12), pp.2407-2418,
377 https://doi.org/10.1016/j.enconman.2010.05.005.

378 Arslan, D. and Özcan, M.M., (2011). Drying of tomato slices: changes in drying kinetics, mineral contents,
379 antioxidant activity and color parameters Secado de rodajas de tomate: cambios en cineticos del secado,
380 contenido en minerales, actividad antioxidante y parametros de color. CyTA-Journal of Food, 9(3),
381 pp.229-236, https://doi.org/10.1080/19476337.2010.522734.

382 Babu, A.K., Kumaresan, G., Raj, V.A.A. and Velraj, R., (2018). Review of leaf drying: Mechanism and
383 influencing parameters, drying methods, nutrient preservation, and mathematical models. Renewable
384 and Sustainable Energy Reviews, 90, pp.536-556, https://doi.org/10.1016/j.rser.2018.04.002.

385 Banout, J., Ehl, P., Havlik, J., Lojka, B., Polesny, Z. and Verner, V., (2011). Design and performance evaluation
386 of a Double-pass solar drier for drying of red chilli (Capsicum annum L.). Solar Energy, 85(3), pp.506-
387 515, https://doi.org/10.1016/j.solener.2010.12.017.

388 Čiplienė, A., Novošinskas, H., Raila, A. and Zvicevičius, E., (2015). Usage of hybrid solar collector system in
389 drying technologies of medical plants. Energy Conversion and Management, 93, pp.399-405,
390 https://doi.org/10.1016/j.enconman.2015.01.051.

391 Chauhan, P.S., Kumar, A. and Nuntadusit, C., (2018). Thermo-environomical and drying kinetics of bitter gourd
392 flakes drying under north wall insulated greenhouse dryer. Solar Energy, 162, pp.205-216,
393 https://doi.org/10.1016/j.solener.2018.01.023.

394 Doymaz, I. (2007). Air-drying characteristics of tomatoes. Journal of Food engineering, 78(4), 1291-1297,
395 https://doi.org/10.1016/j.jfoodeng.2005.12.047.

396 El-Sebaii, A.A. and Shalaby, S.M., (2013). Experimental investigation of an indirect-mode forced convection
397 solar dryer for drying thymus and mint. Energy Conversion and Management, 74, pp.109-116,
398 https://doi.org/10.1016/j.enconman.2013.05.006.

399 Fudholi, A., Sopian, K., Yazdi, M.H., Ruslan, M.H., Gabbasa, M. and Kazem, H.A., (2014). Performance
400 analysis of solar drying system for red chili. Solar Energy, 99, pp.47-54,
401 https://doi.org/10.1016/j.solener.2013.10.019.

402 Fudholi, A., Sopian, K., Bakhtyar, B., Gabbasa, M., Othman, M.Y. and Ruslan, M.H., (2015). Review of solar
403 drying systems with air based solar collectors in Malaysia. Renewable and Sustainable Energy Reviews,
404 51, pp.1191-1204, https://doi.org/10.1016/j.rser.2015.07.026.

405 Jain, D. and Tewari, P., (2015). Performance of indirect through pass natural convective solar crop dryer with
406 phase change thermal energy storage. Renewable Energy, 80, pp.244-250,
407 https://doi.org/10.1016/j.renene.2015.02.012.

Kant, K., Shukla, A., Sharma, A., Kumar, A. and Jain, A., (2016). Thermal energy storage based solar drying systems: A review. Innovative food science & emerging technologies, 34, pp.86-99, https://doi.org/10.1016/j.ifset.2016.01.007.

Khanlari, A., Sözen, A., Şirin, C., Tuncer, A.D. and Gungor, A., (2020). Performance enhancement of a greenhouse dryer: Analysis of a cost-effective alternative solar air heater. Journal of Cleaner Production, 251, pp.119672, https://doi.org/10.1016/j.jclepro.2019.119672.

Lopez, A., Iguaz, A., Esnoz, A. and Virseda, P., (2000). Thin-layer drying behaviour of vegetable wastes from wholesale market. Drying technology, 18(4-5), pp.995-1006, https://doi.org/10.1080/07373930008917749.

Mishra, S., Verma, S., Chowdhury, S. and Dwivedi, G., (2021). Analysis of recent developments in greenhouse dryer on various parameters-a review. Materials Today: Proceedings, 38, pp.371-377, https://doi.org/10.1016/j.matpr.2020.07.429.

Prasad, J. and Vijay, V.K., (2005). Experimental studies on drying of Zingiber officinale, Curcuma longa l. and Tinospora cordifolia in solar-biomass hybrid drier. Renewable Energy, 30(14), pp.2097-2109, https://doi.org/10.1016/j.renene.2005.02.007.

Prakash, O. and Kumar, A., (2014). Environomical analysis and mathematical modelling for tomato flakes drying in a modified greenhouse dryer under active mode. International journal of food engineering, 10(4), pp.669-681, https://doi.org/10.1515/ijfe-2013-0063.

Prakash, O., Laguri, V., Pandey, A., Kumar, A. and Kumar, A., (2016). Review on various modelling techniques for the solar dryers. Renewable and Sustainable Energy Reviews, 62, pp.396-417, https://doi.org/10.1016/j.rser.2016.04.028.

Sacilik, K., Keskin, R. and Elicin, A.K., (2006). Mathematical modelling of solar tunnel drying of thin layer organic tomato. Journal of food Engineering, 73(3), pp.231-238, https://doi.org/10.1016/j.jfoodeng.2005.01.025.

Sethi, V.P. and Dhiman, M., (2020). Design, space optimisation and modelling of solar-cum-biomass hybrid greenhouse crop dryer using flue gas heat transfer pipe network. Solar Energy, 206, pp.120-135, https://doi.org/10.1016/j.solener.2020.06.006.

435 Shalaby, S.M. and Bek, M.A., (2014). Experimental investigation of a novel indirect solar dryer implementing
436 PCM as energy storage medium. Energy conversion and management, 83, pp.1-8,
437 https://doi.org/10.1016/j.enconman.2014.03.043.

438 Sharma, N., Garcha, S. and Singh, S., (2019). Potential of Lactococcus lactis subsp. lactis MTCC 3041 as a bio-
439 preservative. Journal of Microbiology, Biotechnology and Food Sciences, 2019, pp.168-171.

440 Singh, V., Hedayetullah, M., Zaman, P. and Meher, J., (2014). Post-harvest technology of fruits and vegetables:
441 An overview. Journal of Postharvest Technology, 2(2), pp.124-135.

442 Singh, S. and Kumar, S., (2012). Testing method for thermal performance based rating of various solar dryer
443 designs. Solar Energy, 86(1), pp.87-98, https://doi.org/10.1016/j.solener.2011.09.009.

444 Suárez, M.H., Rodríguez, E.R. and Romero, C.D., (2007). Mineral and trace element concentrations in cultivars
445 of tomatoes. Food Chemistry, 104(2), pp.489-499, https://doi.org/10.1016/j.foodchem.2006.11.072.

446 Taiwo, A. and Bart-Plange, A., (2016). Factors responsible for post-harvest losses and their effects on rice
447 producing farmers: A case study of Afife and Aveyime rice projects in the Volta region of Ghana.
448 International Research Journal of Engineering and Technology, 3(4), pp.2395-0056.

449 Vijayan, S., Arjunan, T.V. and Kumar, A., (2020). Exergo-environmental analysis of an indirect forced
450 convection solar dryer for drying bitter gourd slices. Renewable Energy, 146, pp.2210-2223,
451 https://doi.org/10.1016/j.renene.2019.08.066.

452 Wang, J., Xiong, Y.S. and Yu, Y., (2004). Microwave drying characteristics of potato and the effect of different
453 microwave powers on the dried quality of potato. European Food Research and Technology, 219(5),
454 pp.500-506, https://doi.org/10.1007/s00217-004-0979-1

Figures

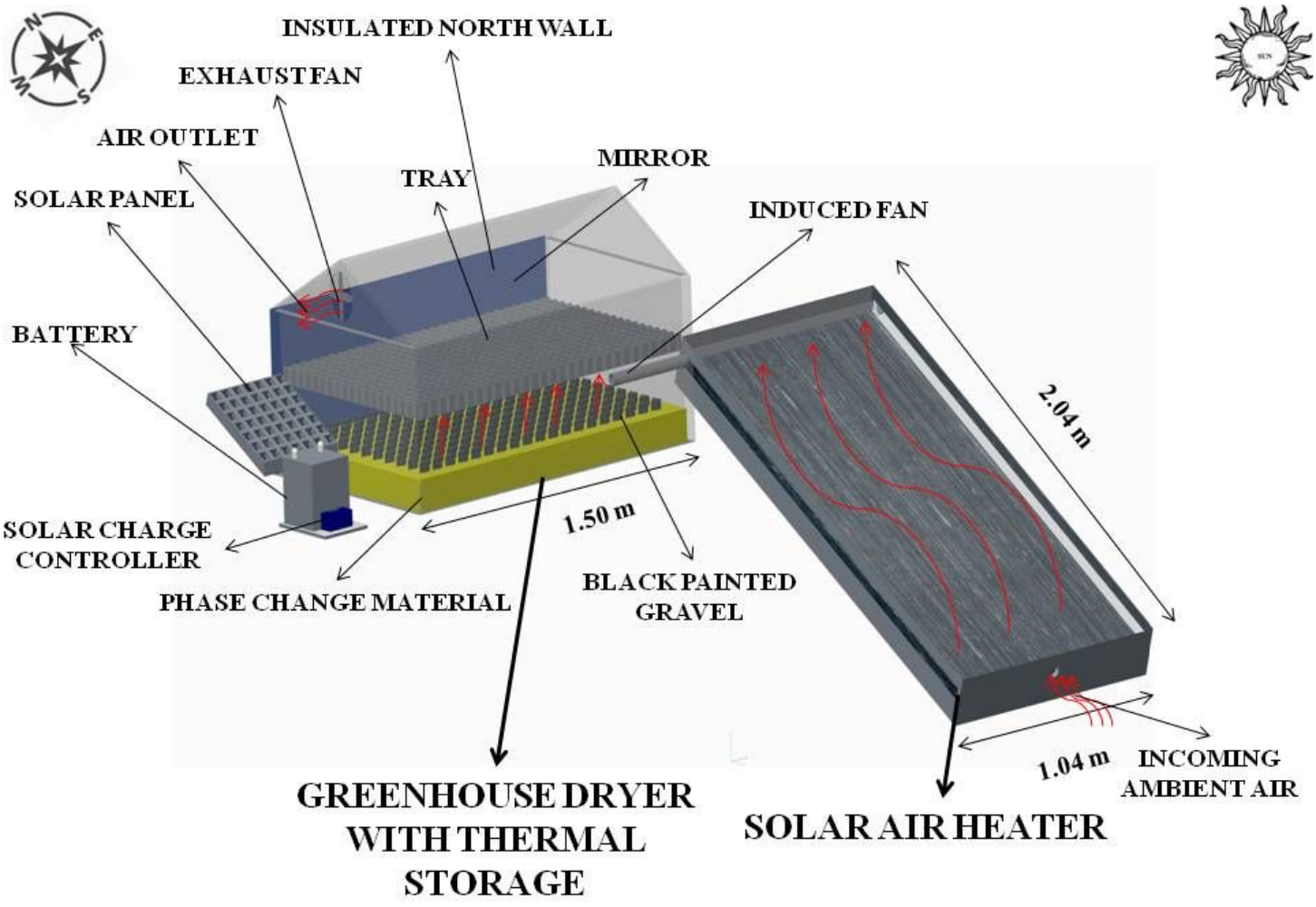


Figure 1

Schematic Diagram of Hybrid Greenhouse Dryer

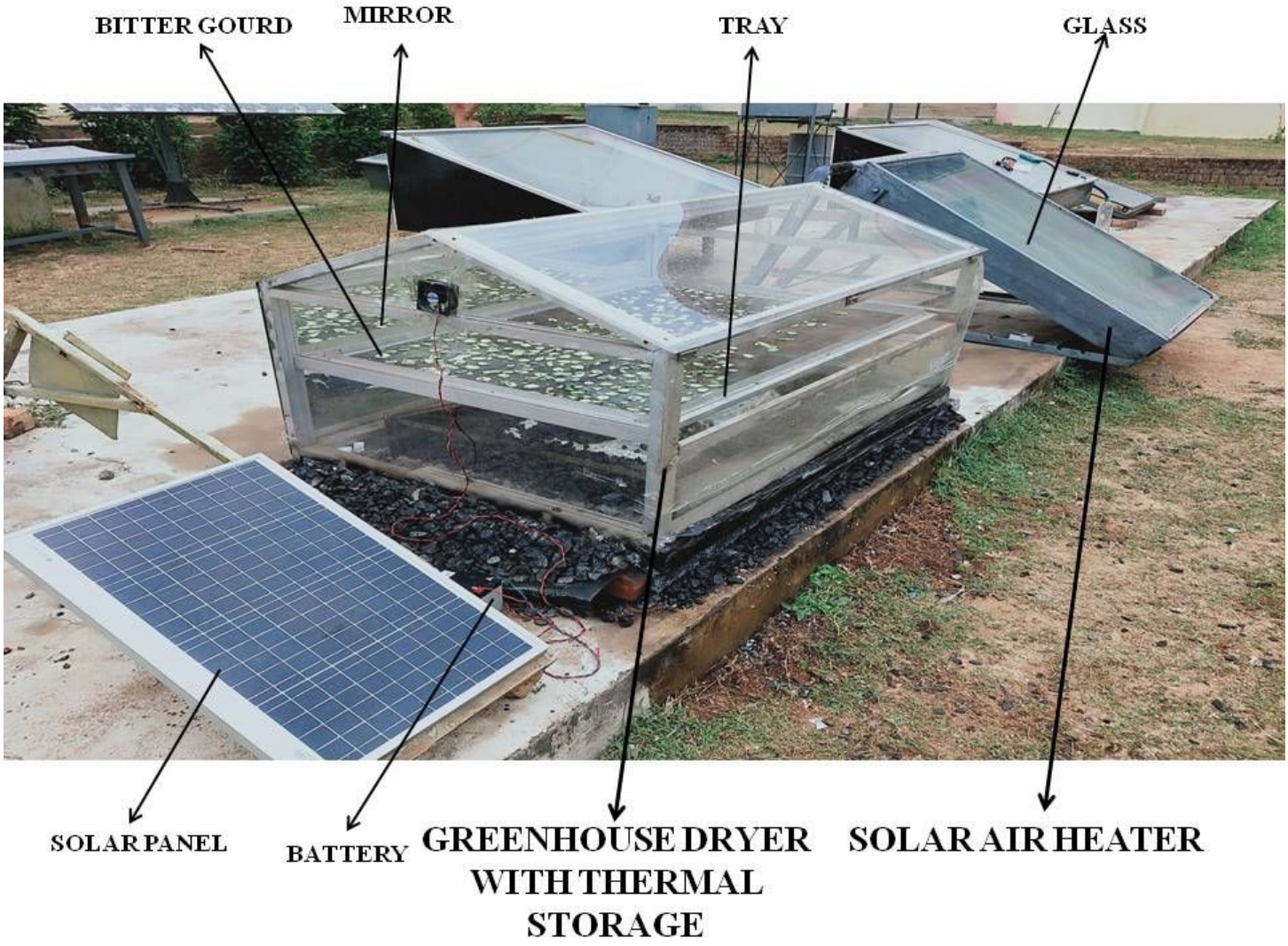


Figure 2

Experimental Setup of Hybrid Greenhouse Dryer

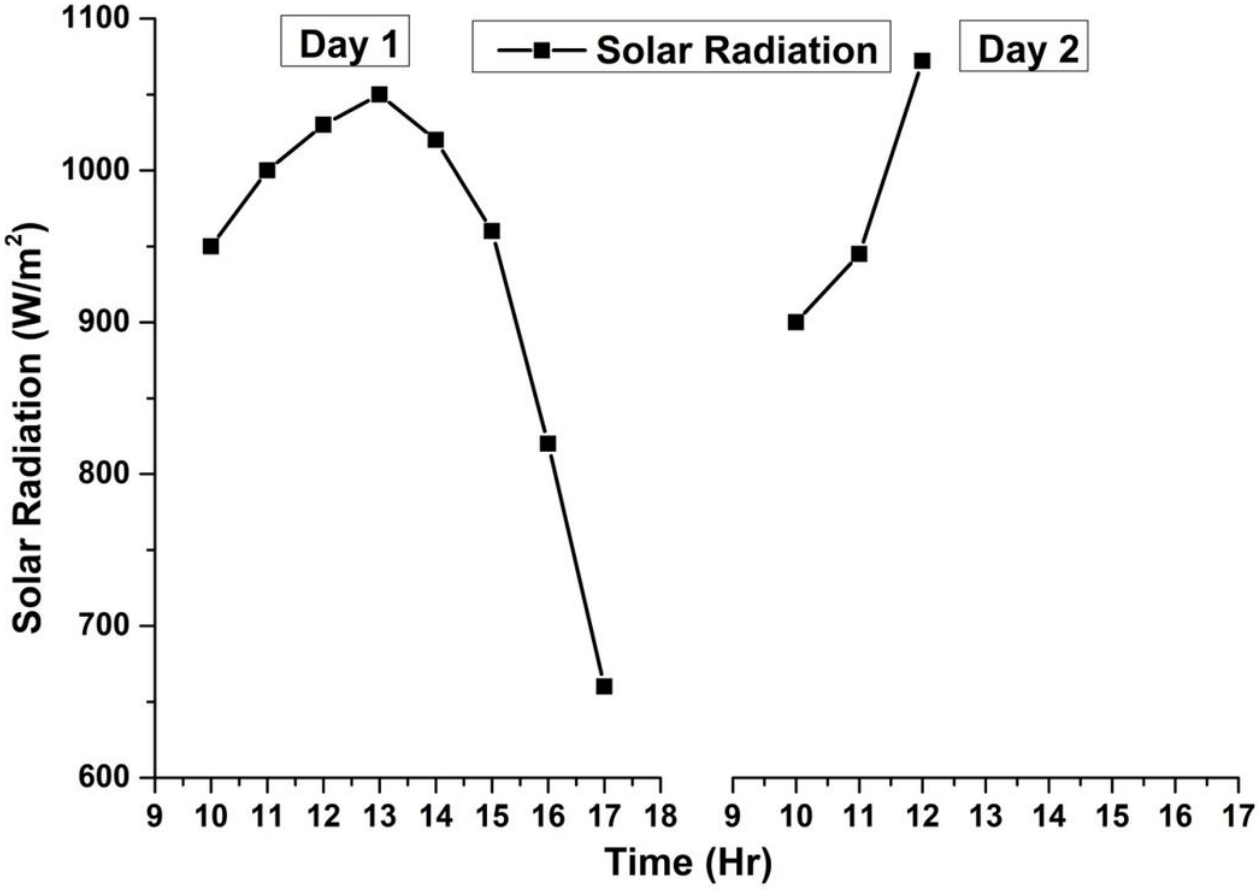


Figure 3

Variation of Solar Radiation in Day 1 & Day 2

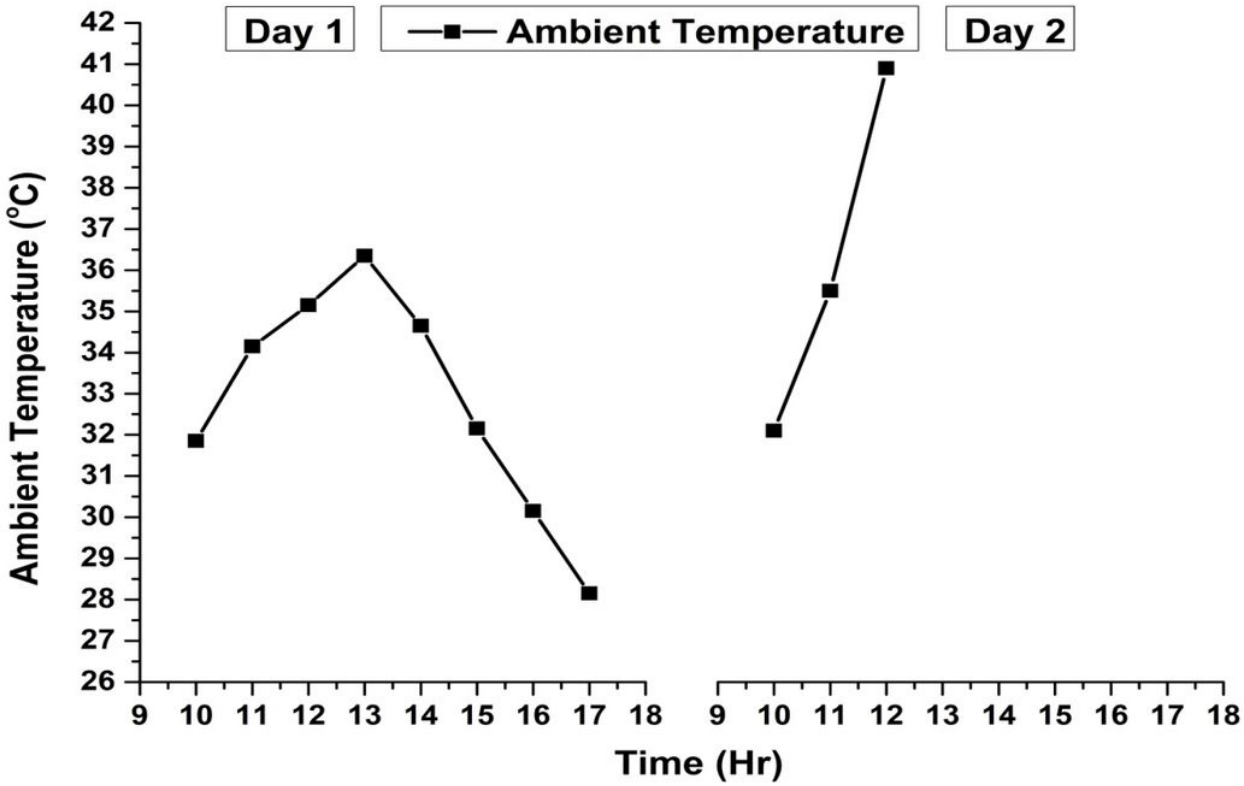


Figure 4

Variation of Ambient Temperature in Day 1 & Day 2

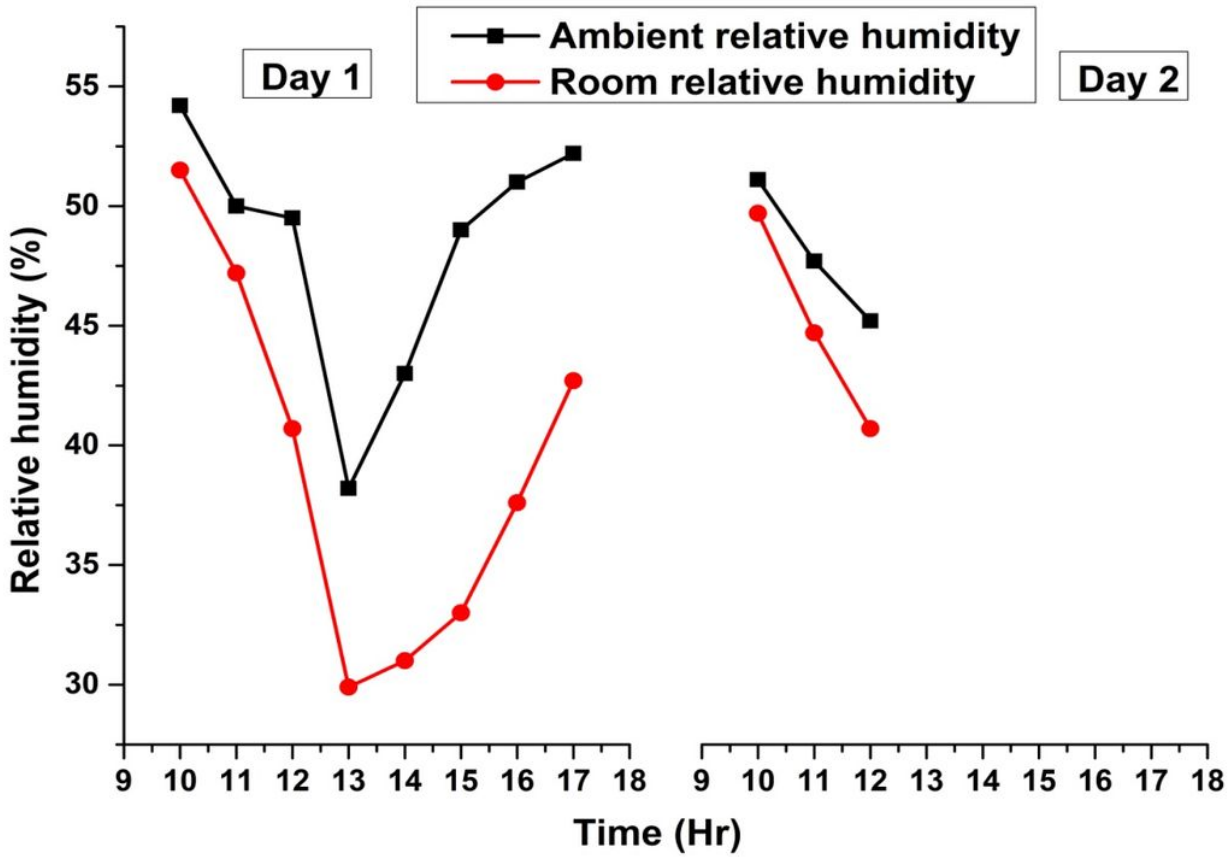


Figure 5

Variation of Relative humidity in Day 1 & Day 2

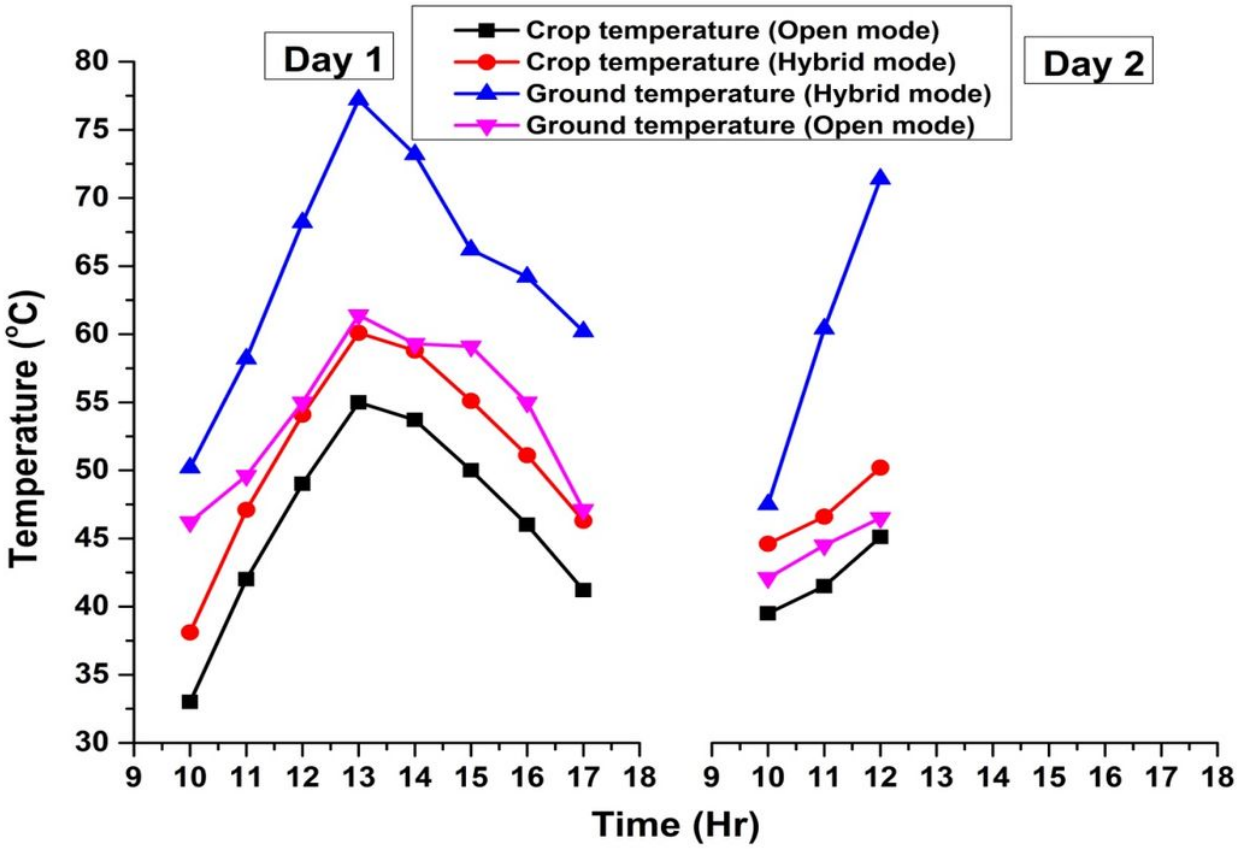


Figure 6

Variation of crop & room temperature in Day 1 & Day 2 for Hybrid and Open Mode

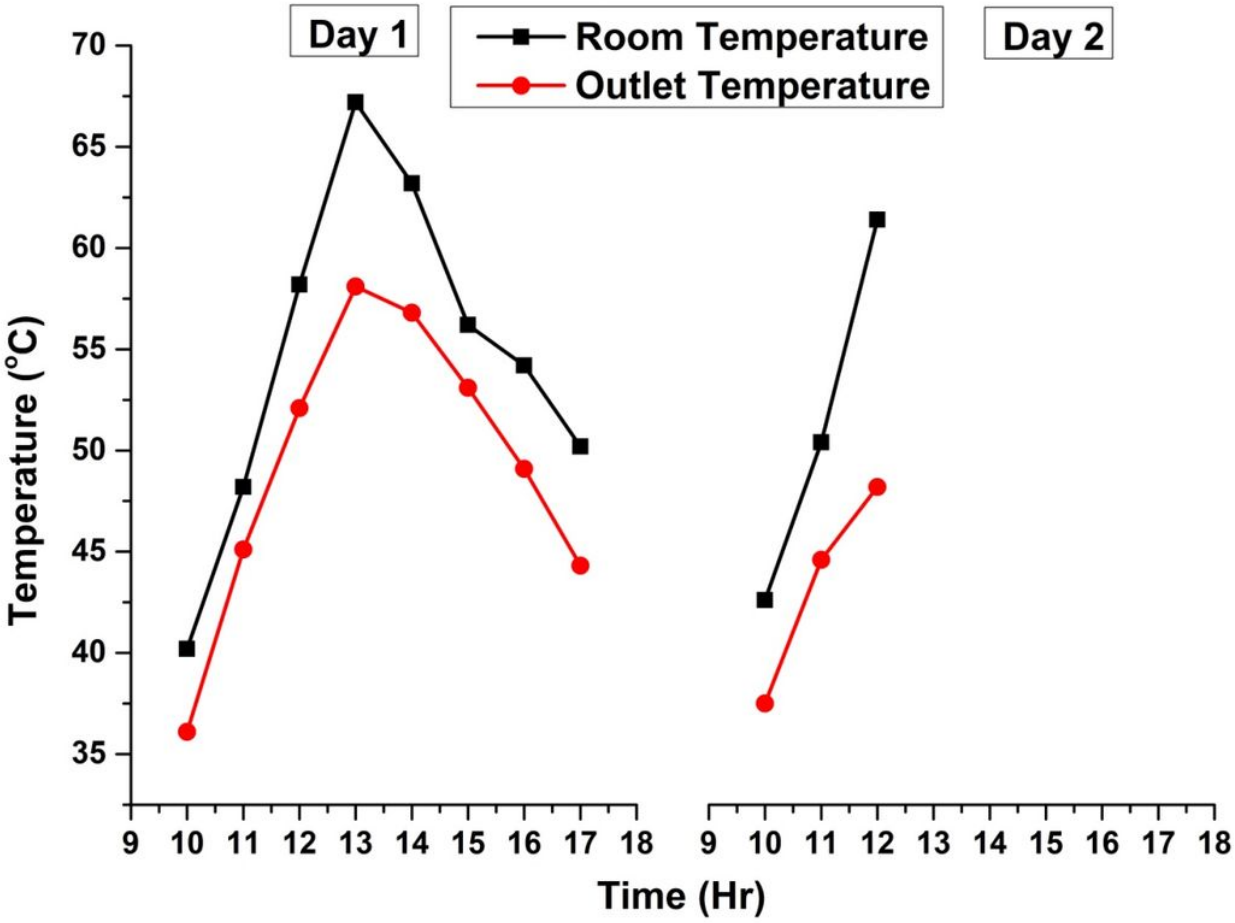


Figure 7

Variation of Room Temperature & Outlet Temperature in Day 1 & Day 2

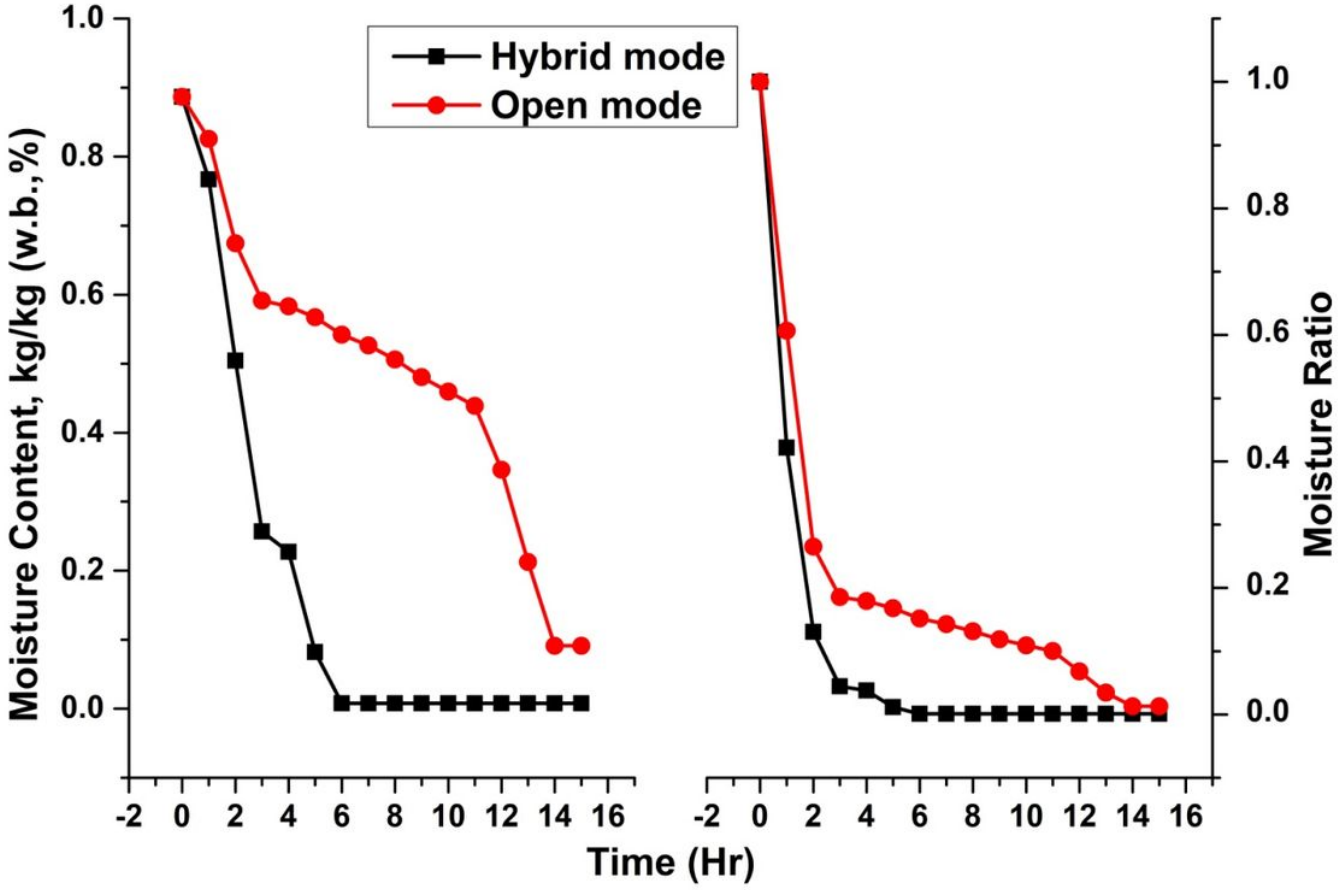


Figure 8

Variation of Moisture Content & Moisture Ration for Hybrid & Open Mode

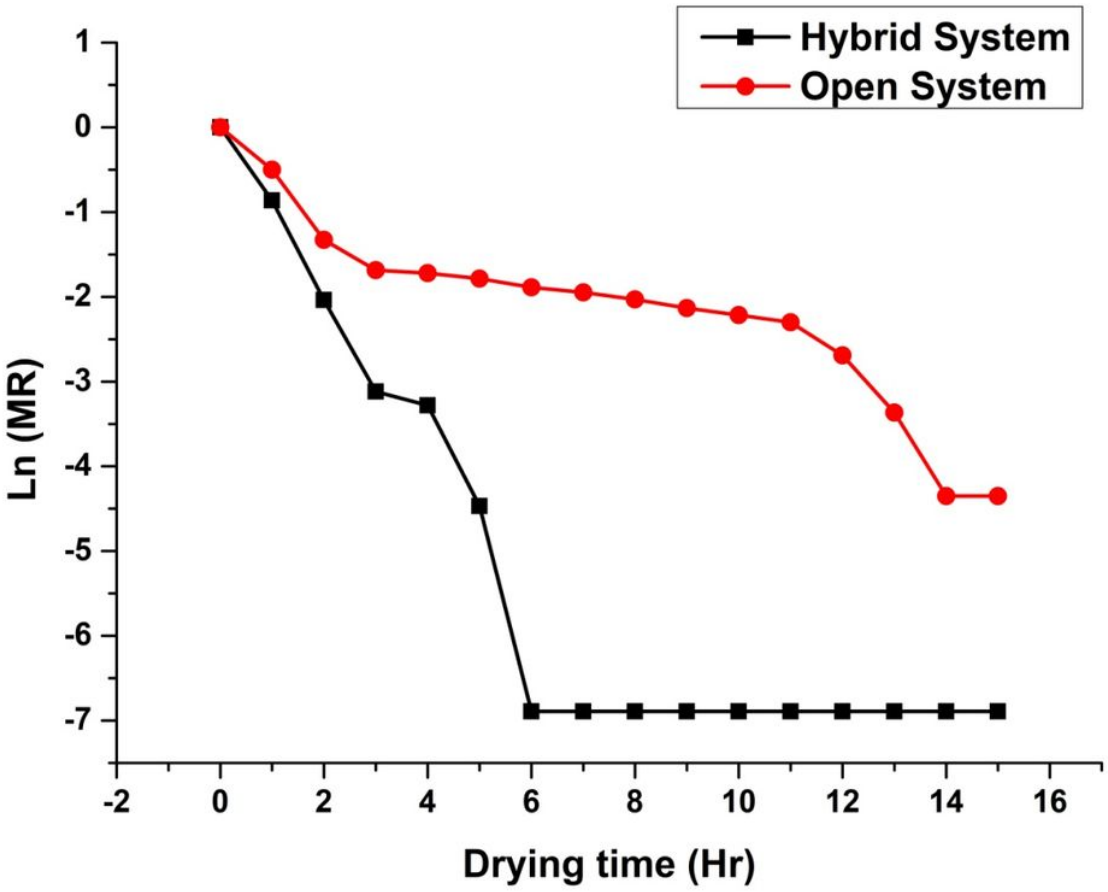


Figure 9

Variation of Ln MR and drying time for drying of bitter gourd flakes

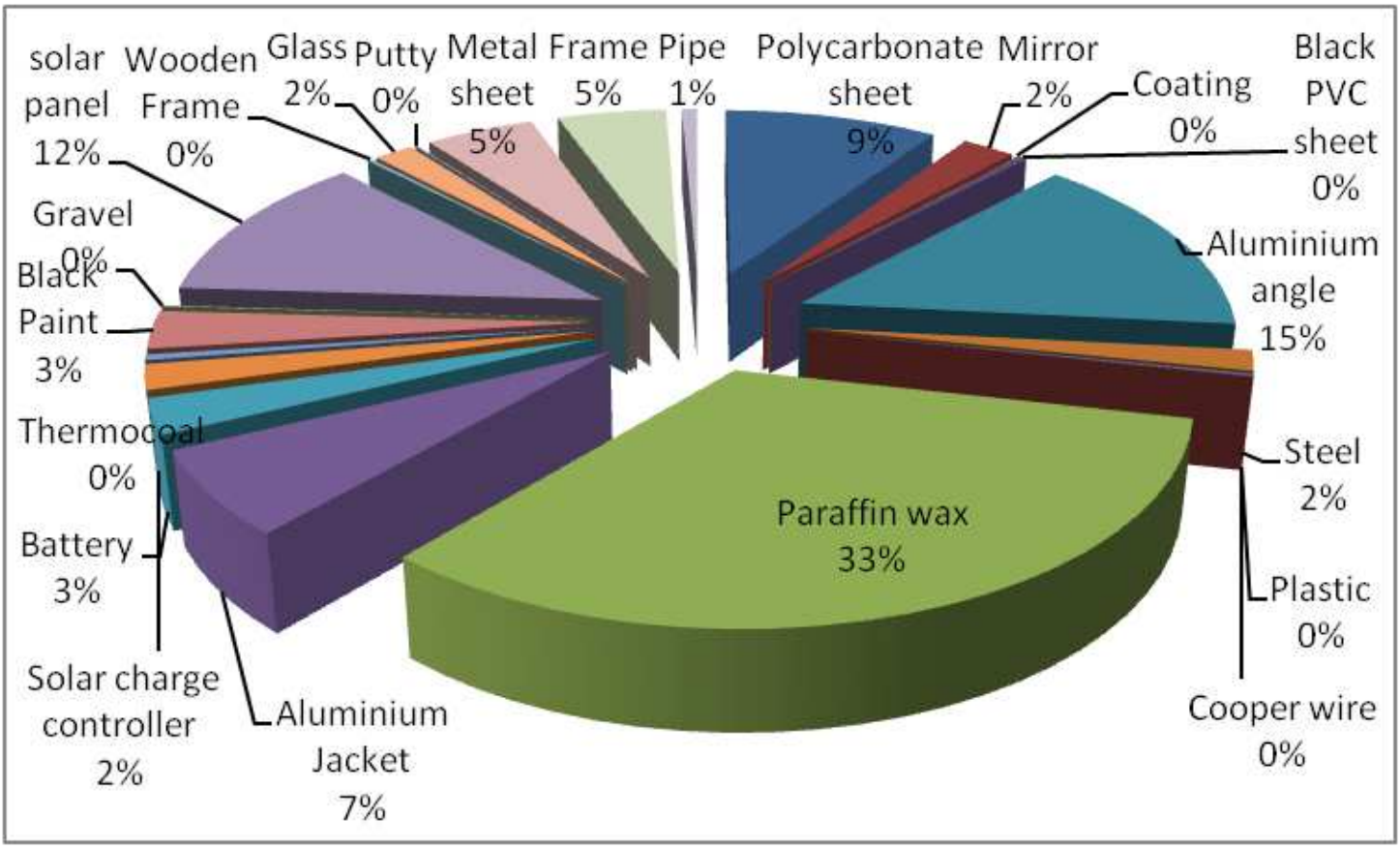


Figure 10

Breakup of Embodied Energy for Hybrid Greenhouse Dryer

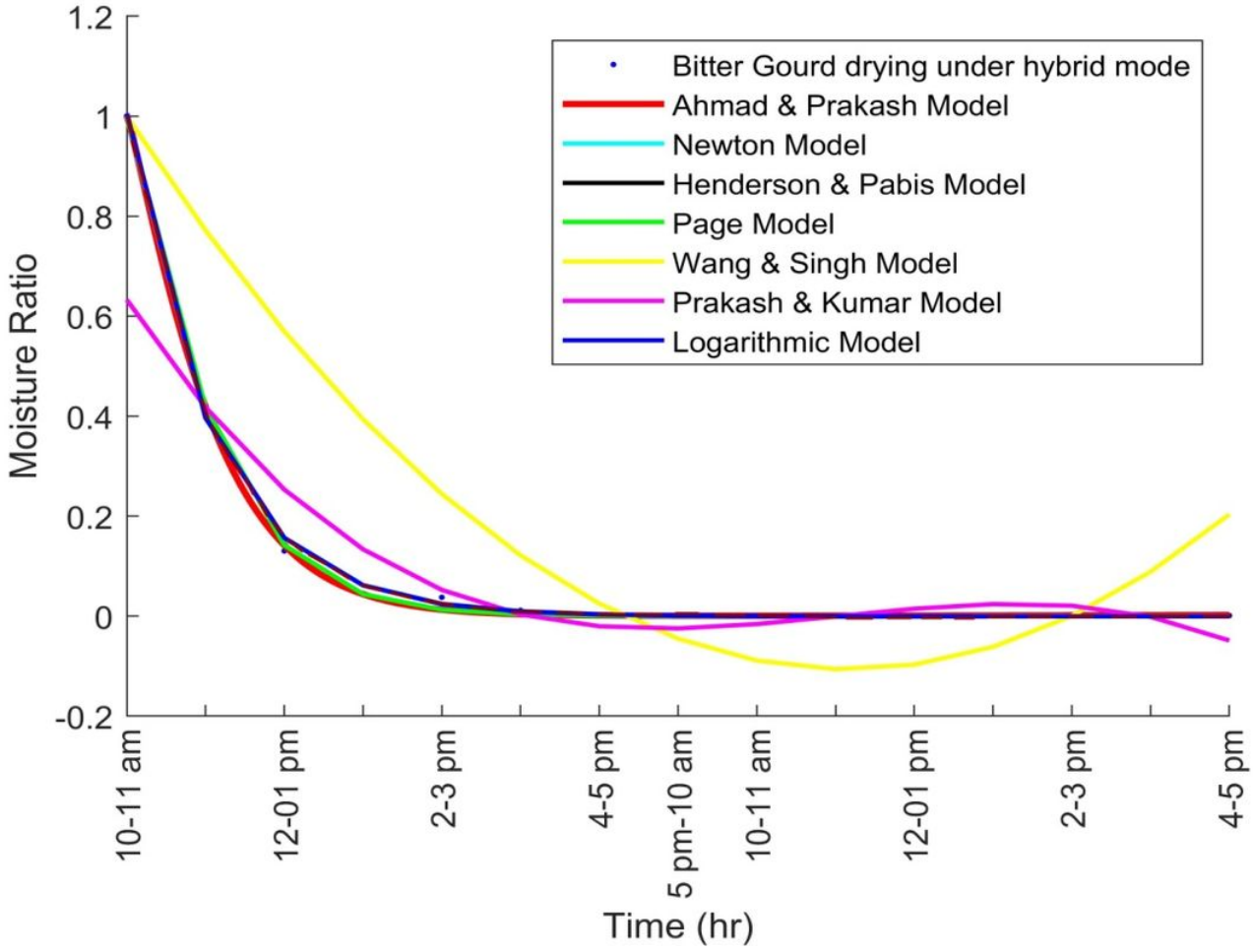


Figure 11

Variation of Prediction & Experimental Value of Moisture ration for different Mathematical Model in Hybrid Mode

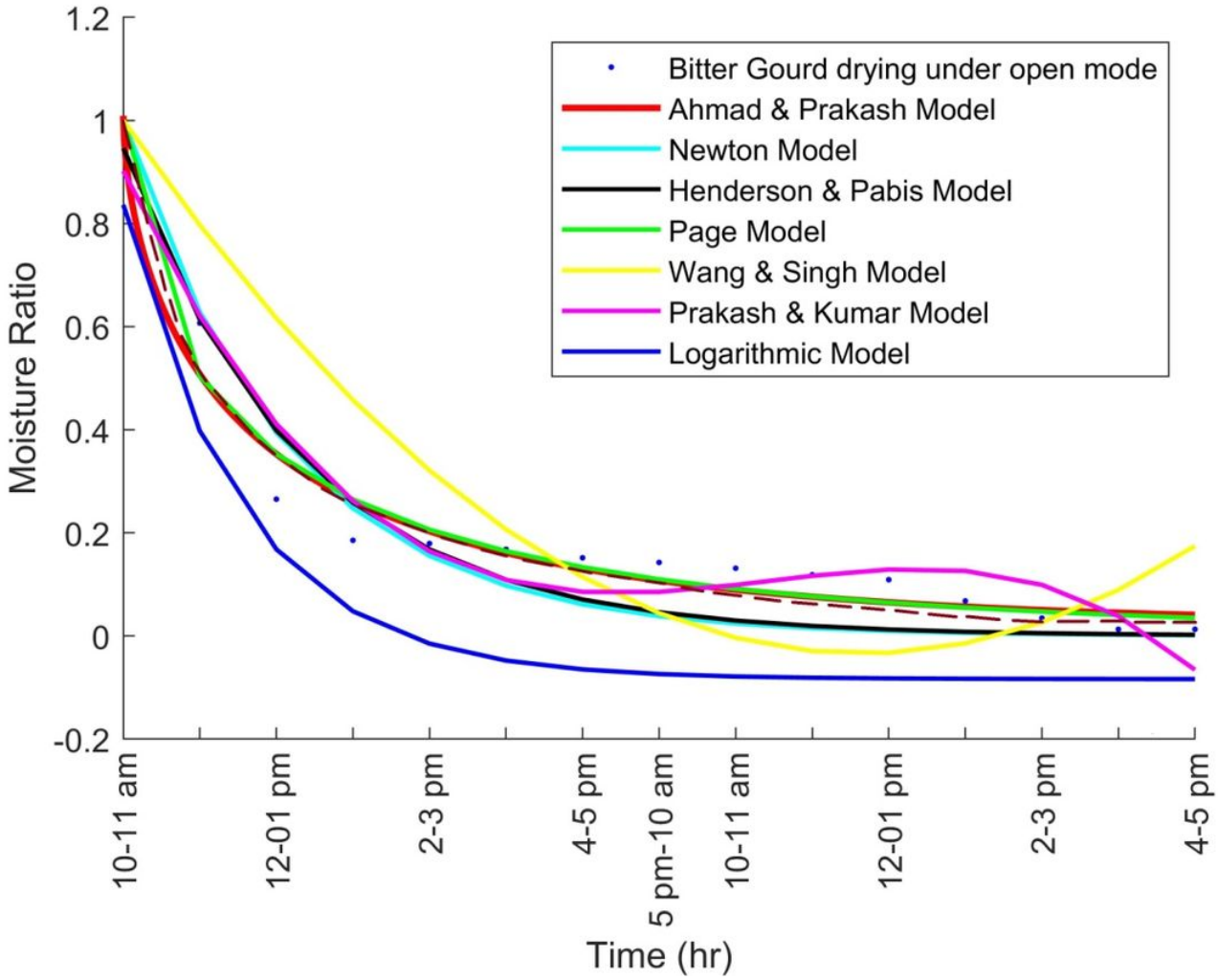


Figure 12

Variation of Prediction & Experimental Value of Moisture ration for different Mathematical Model in Open Mode

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- BittergourdTable.pdf

International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. #, No.# (Mon-20..)

Dynamic Visualization and Analysis of Government Responses-A Support System to Control Pandemic Situations

Neelu Chaudhary[1], Hardeo Kumar Thakur[2], and Rinky Dwivedi[3]

[1] *CST, Manav Rachna University, Faridabad, India*
[2] *CST, Manav Rachna University, Faridabad, India*
[3] *CSE, Maharaja Surajmal Institute of Technology, Delhi, India*

E-mail address: neelu@mru.edu.in, hkthakur@mru.edu.in, rinkydwivedi@msit.in

Abstract: Covid 19 has focused the world's attention on health care facilities across the globe. Data modeling is a latent contributor for dynamic assessment and reporting/analyzing in distributed research networks. During our research, it was found that without demonstrative data, data sharing among subgroups focusing on various policy-making classifications, discoveries, and predictions become problematic and inefficient. The dearth of illustrative pandemic data (COVID-19) is a holdup for showcase and deliberate potential solutions. The realization of data for pioneering new technology in clinical healthcare systems and designing government decision support systems to help fight novel coronavirus is crucial. Dynamic graphs are a potential tool that aids in the dynamic assessment of country-wise policy data at a given point of time and is extremely significant for estimating public health measures and stringent policies. In the present research, the openly accessible Oxford's COVID-19 Government Response Tracker (OxCGRT) dataset is used. The OxCGRT is counting statistics from more than 180 countries, it enables the investigators and officials to explore the pragmatic outcomes of policy responses on the accelerating spread of COVID-19, in addition to financial and communal welfare. The stringent policy index and Containment and health policy index is observed for the 11 Countries across the world with the highest occurrence of cases. This assortment of the sub-group countries reduces the complexity and upsurges the representativeness of data with respect to the model policy indicators.

Keywords: Dynamic Graph, Covid-19, Government Responses, Policy Indicator, Social Network.

1. INTRODUCTION

A dynamic graph is a method of representing time series network which represents vertex as entity and an edge between two vertex represents relation between them. The weight of edges represents strength of connection. The entities and their connections are dynamic i.e., varies with time. Using the dynamic network approach, the study explores the characteristics of the network structure for the countries round the globe and their relative policy responses in retort to the COVID 19[1,2]. The dataset demonstrated in the present research is extracted from the Oxford's COVID-19 Government Response Tracker (OxCGRT). OxCGRT collects the statistics through a sequence of four policy indices, namely-Overall government response index • Stringency index • Containment and health index • Economic support. These indices allow the competent and unpretentious across realm assessment of administration involvement in handling the pandemic situation. The current research explicitly emphasis on two extents of OxCGRT dataset – one is stringency index, and another is Containment and health index. Oxford University establishes, the "Stringency Index" that computes the actions instigated by the authorities and administrations across the globe to tackle the current challenging condition. Whereas, the "Containment and Health index" computes the lockdown restrictions and specific health trials in the process of testing policy, investment in vaccine etc. treaded up by the authorities.

The knowledge behind the assortment of these two indices – as the world is facing an economy downtown, Covid 19 has focused everyone's attention on healthcare and investors' attention on companies in the health sector. Different sub-sectors are experiencing different impacts-for example, there is a major increase in drug call as patients are more concerned about the health as well as the demand is anticipated to increase in future also [3,4]. How authorities will adapt these healthcare challenges and forthcoming difficulties will determine how rapidly countries economy will get recovered.

E-mail:author's email

This graphical design is intended to persuade added exhaustive study [5] and to reveal the potential for legislators, scholars, inventors and investors to address significant public policies and epidemiological queries using OxCGRT data. The study epitomizes a government response network that assists in understanding the discrepancy consequences of policy indices on government response index specifically on - health sector parameters. The designed network comprises of nodes (countries and policies) and edges (indices value of each policy), the premediated network is further visualized and analyzed through the support of standard Gephi software. The contribution of the present research work can be summarized as: -

The research is among the first to present a visual dynamic approach to assess the two major dimensions-stringency factors and health and containment factors across the globe for COVID-19 pandemic. However, paper further presents a comparative analysis of customized government policies towards the specific nations like-China, Indonesia, Malaysia etc. The prime focus of the research is to evaluate the common factors that are shared by majority of nations and need to be addressed in the pandemic situation. The usage of dynamic graphs provides a visual dynamic approach that further helps to develop more precautionary measures, cures and solution strategies.

Primarily, it starts with the Visualization of disparity in government response network in accordance to the time variances. Further the research moves towards the comprehensive Analysis of the shaped network grounded on the policies indices and government indexes. The subsequent vital step is to identify the patterns that helps the effect of vaccination policy in refining the government response index. The paper is organized as follows: Section 2 consists of literature survey that covers the background and related work of the considered domain. It also covers the comparative analysis of relevant work to present research. Based on which, the main objectives behind this work are also specified in this section, that are well formulated in form of questions. Section 3 describes the workflow of this study which addresses the data elicitation part, covers pre-processing and conversion of data to be observed that forms the prerequisite for the process, enlightens the network formation and generation of time series network respectively. This section also covers visualization and analysis of dynamic network for conceptualizing the dynamic data and scrutinizing of the network in response to the addressed questions. In section 4, the attained results are presented and discussed established on the objectives of the study. Finally, section 5 covers the conclusion that summarize the discussion.

2. LITERATURE SURVEY

World is fighting with two battles simultaneously; one is against pandemic or health sector deficiencies and other is against the slowdown of world's economy. Innovative reforms need to be ensured in healthcare sector like consulting doctors across the country. History shows, innovation thrives during time of crises - take the financial crises of 2008 which led to the wide spread option of cloud computing. The cloud happens around early 2000 but it gained a new footing as the economic slowdown took place. As well as accelerating the adoption of technologies, crises can also foster the development of new ideas like at the time of second world war, digital computers, jet engines etc. are invented nuclear technology for nuclear weapons.

The pandemic is accelerating, as of mid-May 2021, the novel coronavirus disease (COVID-19) has triggered the contamination of about 160,813,869 people and accounts for more than 3,339,002 deaths (WHO, 2021).

A stringent lockdown is consigned by many countries. In the primary stage of covid-19 outbreak, countries like India, China, France, Italy etc. had imposed the strictest lockdown that is further relaxed outside the red zones. As per the data, Spain imposed strictest measures much later in its case and death count than all others. Sweden maintained the most liberal measures in this set, and Iran the second most liberal [6].

Although lockdown is an immediate solution, but the livelihood of the family's needs to be safeguard, there must be a return to normality. Due to the financial crises, few states and authorities abandon policies like ban on transportation, closing business and mass confinement [7,8].

In 1980s when neoliberalism, spread across the world private enterprises and open market became economic buzzwords while government reduced their roles as regulators this implies the major countries across the world in health care sector plays as per the free market rules, relying highly on Intellectual Property Rights (IPR) to guard their products. Lifesaving drugs are expensive and because of IPRs their formulations could never be shared. This becomes a bottle neck for multilateral organizations like WHO to operate and formulate a global plan to handle the current pandemic situation.

In 2020, when pandemic hits the world WHO created a COVID-19 technology access pool (C-TAP) to share knowledge with manufacturers of developing countries, but no companies signed for it they simply don't want to violate IPR.Exactly how many years it needs to nullify or handle the current pandemic situation entirely, depends upon the behaviour of the people and on the policies formulated by the countries across the world. During the research, it was discovered that there is a scarcity of global proposals for handling the pandemic situation. Many nations specific proposals were found stating their explicit complications along with the measures that their government took to handle the situation. Table 1 presents

the comparative analysis of some relevant to the present research.

TABLE I. COMPARATIVE ANALYSIS WITH RELEVANT RESEARCH

Countries	Purpose	Methodology	Findings	Research limitations
China [9,10]	To state the association amongst trust and preventive measures taken by government during the pandemic.	Data from 3000 Chinese adults conducted using quota-sampling method.	The constructive association between Government trust and excessive precautionary behaviour was found to be statistically substantial.	Trust is the only emotion analysed among the people with less knowledge and people with less negative emotion for evaluating the preventive behaviour during COVID 19.
Indonesia [11]	Etymological behaviour of the government is analysed to improve the confidence and self-esteem of the citizens	Positive Discourse Analysis (PDA) is reformed for communal variations through resistance discourse	Authorities made use of submission, predication, argumentation and mitigation policies to improve community 's optimism	Not talks about the government strategies at root level.
Malaysia [12]	To induce the familiarity levels, attitudes and practices toward pandemic midst the Malaysian public.	a quantifiable, cross-sectional survey was estimated for the Malaysian background. Online data gathering using the Inspection Monkey platform.	Efficacious control of COVID-19 is found (83.1%), the way the Malaysian government was handling the crisis (89.9%).	The cross-sectional online survey was conducted for 27th March and 3rd April 2020 only
Ethiopia [13,14]	Evince data on the - Public knowledge, Attitude, and Practice (KAP) is gathered to testify the response of the service providers in Addis Ababa, Ethiopia	self-govern opinion poll and observational taxation using a specification. For statistical analysis SPSS version Evocative information, association quantity and chi-square tests were performed.	More than 60.7% of the contributors had a positive attitude towards preventive measures by authorities. An optimistic correlation between knowledge and attitude is found.	Severity policies are explored. However, inhibition and healthiness index are yet to be revealed
Yogyakarta [15]	To reconnoiter the grounding of government policies towards the revitalization of economy in the form of Micro, Small and Medium Enterprises (SMEs) after Corona Virus Disease	A expressive investigation method of principal and subordinate data attained from field survey and publication sources	Business Model Canvas (BMC) becomes an applied and apposite substitute for economic reform.	This study is focused on stringency factors like - social distancing and physical distancing
Macao [16]	How government or policymakers might lessen the harmful impressions of pandemic and to classify operative policies for crisis management.	Both qualitative exploration and archival research method is used for eliciting primary data from the stakeholders and secondary data on the pandemic facts	crisis administration approaches to lessen devastation were organized based on - suitable procedures, opportune choices and the role of the government.	Sample size is mainly focussed on the people associated with gaming and financial institutions only

Therefore, to get an insight on how the policy indicators varied with respect to time and how these variation affects the government responses, we analyze the dynamic network based on questions as follows:

- RQ-1-What will be the effect of disparity in policy indices on the country wise stringency index?
- RQ-2-What will be the effect on country wise containment health with respect to these policy Indices?
- RQ-3-Significance of each policy and what will be their effect on the prevention of complete lockdown?
- RQ-4-Effect of vaccination policy in improving government responses to Covid-19?

A. Why dynamic graphs are used??

From decades, modelling is extensively used in defining the finest approaches or policies for the moderation of the influence of infectious viruses. Presently, the demonstrating of a composite structure for instance the outburst of COVID-19 contagion is amid the contemporary issues. The upsurge of the infectious diseases is at the prime attention of the world municipal and plays a critical role in decision making [17,18]in almost every sector i.e. public health sector, innovation sectors, industrial sectors or educational sectors etc. Apart from all disaster created by Covid -19, There is a less familiar story that the crises are boosting innovation with the emergence of brand-new ideas like mRNA vaccine, drone deliveries a new technology and with application of old ones with surprising new ways. The crises have turn out to be an acceleration of adopting technologies and new lessons to be learn for post pandemic era.

By representing the static tabular data sets in the form of dynamic graphs [19,20,21,22], more precautionary measures, cures and solution strategies may be discovered. Research around the world may get pioneered to advance new procedures or practices for the development of vaccines, drugs and medical equipment's etc. in months rather than in years. The aim of this artefact is to organize and formally represent the dynamic knowledge base [21] of novel corona virus in the form of dynamic graphs [22,23] based on the government policies towards the COVID-19 infection spread. This implies, at the first stage, understanding of the – government policies, their social and financial aftermaths, implementation issues and apprehensions that persons partake and the possible set of dynamic solution strategies [24]. Since most data is recorded in the form of text data or tabular data measuring the impact of policies is the vital chunk of information influencing the of the situation of people during pandemic. This is where dynamic graph modelling can play a crucial role.

3. WORKFLOW

The present research is inspired by the recent attempts to apply network analysis to researchers' community. The work consists of visualizing the co-policy network of observed OxCGRT dataset on response indexes. "Fig 1" depicts the workflow diagram of the process initiating from data collection to processing of data, further moves towards the generation of time series network and finally at end visualization and analysis of the designed network corresponding to the considered dataset is shown. The importance of analyzing the co-policy network for finding the results or answers lies in the ease to visualize as compared to the textual representation. Network analysis is used to understand aggregate patterns of connections and interactions of inter government policies and intra government policies towards the novel corona virus. The inter- actions among policies and countries helps to visualize effect on government responses i.e. stringency and containment health with respect to time. In the current research, the dynamic network visualization and analysis tool namely Gephi is used. The Gephi network visualization tool helps in analyzing the complex dynamic network [25,26,27] and is also empowered to demonstrate statistical analysis like degree distribution, modularity and cluster efficient.

A. Data Collection

For the analyses purpose, Oxford COVID-19 Government Response Tracker (OxCGRT) is customized, that reports the prerequisite for uninterruptedly updating, readily functioning and comparable statistics on policy measure. OxCGRT intelligences information on 20 indicators of government response of 186 nations across the globe. The indicators are of three types namely-*ordinal, numeric and text*. Ordinal pointers are used to quantify policies on a modest gage of sternness or intensity. Although, daily testification of data by the indicators is available, but to preserve the simplicity and clarity of process in our research we concentrate on fortnight data. The aggregated information of four policy indices as: Overall government response index, Stringency index, Containment and health index and Economic support index

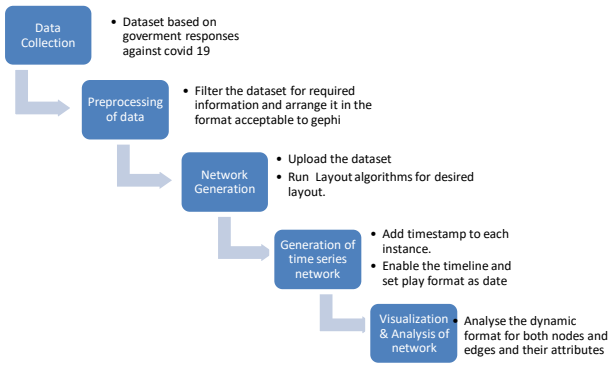


Figure 1. Workflow diagram of the designed Time Series Network

is accessible in the dataset for readily use. As mentioned above, the present research concentrates on two dimensions - stringency indices and containment and health indices and shows variation in these Reponses as value of policy indicator changes over time. Each index is composed of a series of individual policy response indicators. For each indicator, a score is created by taking the ordinal value and subtracting an extra half- point if the policy is general rather than targeted [20]. These are further escalated by their maximum value to create a score between 0 and 100, with a missing value contributing 0. The found scores are additionally averaged to get the composite indices. Table IV shows the relation between policy indicators and response indexes. Both indexes are simple averages of the individual component indicators. This is described in equation 1 below where *n* is the number of component indicators in an index and *Ij* is the sub-index score for an individual indicator.

B. Preprocessing of data

For simplification purpose, 11 countries are selected (that may have maximum impact of Covid-19 virus) and the five most influential ordinal policy indicators - Restriction on gatherings, international travel ban, testing policy, contact tracing and Vaccination policy to be observed. However, the model can be extended for numerous other countries and their corresponding policy indicators. Table II. illustrates the policy factor/indicators affecting the calculation of respective Index and their significance.

TABLE II. THE POLICY INDICATORS

Policy	Significance
Restriction Gathering	It accounts the size of - constraint on gatherings. Its value is measured in ordinal scale and binary for geographic scope.
Contact Tracing	It postulates the government effort in contact tracing and if yes then up to which level. Its value is also measured in ordinal scale.
International Travel Ban	It records the limitations on Worldwide travel and is measured in ordinal scale.
Testing Policy	It specifies the measures any government employs to select person/s that can be tested. Its value is also measured in ordinal scale.
Vaccination Policy	It records the principles government trails to select clusters for vaccination and its distribution. It is also measured in ordinal scale.

Table III. exemplifies the comprehensive set of value description corresponding to all the observed policy indices described in Table II. Policies having indices values in ordinal scale and as the scale increases the intensity of policy implementation also increased. Policy indices of policies like restriction on gathering, international travel ban and testing are measured up to scale 4, while contact tracing is measured up to scale 3 and vaccination policy is measured up to scale 5. No data is mentioned in columns where no information is available

The calculated index value is a linear function dependent on the policy indicators or policy factors of concern [28]. Table IV. shows the relationship between the observed Government response Indices and policy indicators. The index value and sub index value of the policy indicators will be calculated from equation (1) and (2) respectively [28]. These equations represent how index and sub-indexes values are calculated and are used for precision purpose. The precise values from dataset are selected and calculated on all policy indicators. The results demonstrate that these values vary in linear manner – that explains the statistic that limited policy indicators show the same variation in index values.

TABLE III. COMPREHENSIVE SET OF VALUE DESCRIPTION OF EACH POLICY INDICATOR VALUE.

Policy Indices	0	1	2	3	4	5
Restriction on Gatherings	None	Restriction on large gathering.	Restriction on 100+ people	Restrictions on 11-100 people	Restrictions on <=10 people	NA
Contact Tracing	No data	No contact tracing	Limited	Comprehensive	NA	NA
International Travel ban	None	Screening	Quarantine after arrival from high-risk regions	Ban on arrivals from some regions	Total border closure	NA
Testing	No data	No testing policy	Those having symptoms/ contact with infected/emigrates/key worker	Anyone showing symptoms	Open public testing	NA
Vaccination	No data	No availability	Any one of key workers/ clinically vulnerable groups / elderly groups	Any two	For all three	Partial addition of groups/ages

For Calculating Index Value

$$Index = \frac{1}{n}\sum_{j=1}^{n} I_j \quad (1)$$

Where: n- is the number of policy indicators. I-Index value calculated for each indicator.

For Calculating Sub-Index Value

$$I_{j,t} = 100\ (P_{j,t} - 0.5(F_j - f_{j,t}))/M_j \quad (2)$$

M_j is the maximum value of the indicator.

Flag (F_j =1 if the indicator has a flag variable, or 0 if the indicator does not have a flag variable.

$P_{j,t}$ is the recorded policy value on the ordinal scale

flag ($f_{j,t}$) =1 if the recorded binary flag for that indicator is true.

TABLE IV. RELATIONSHIP BETWEEN GOVERNMENT RESPONSE INDEX AND POLICY INDICATORS.

Response Indexes ⟶ / Policies ↓	Stringency index	Containment and health index
Restriction on gatherings	√	√
International Travel Ban	√	√
Testing Policy	-	√
Contact Tracing	√	√
Vaccination Policy	No data	No data

C. Network generation for visualization and analysis

Using Gephi, the visualization and analyzing tool of all types a network, the time series data is converted in form of nodes and edges. This process requires a thorough knowledge of the tool internal algorithms since the prerequisites required for the process needs to be arranged in a proper format supported by the visualization and analysis implementation. To maintain the consistency of the process separate sheets for nodes and edges must be created. A major part is to stipulate the timestamp or quantum (for dynamic graphs) for a node or edge to be present in the graph and decision support system to envisage its attribute value at that timestamp.

D. Generation of time series network

The government response network for the visualization and analysis purpose have separate nodes for distinct countries and policy indicators. Table V demonstrates the statistics about the network shaped in Gephi tool. The edges show the indices value of each policy employ by individual countries at any point of period. The government response index value has been taken as attribute of each country and the effect of variation in value of policy indicators over response index is illustrated. By converting a textual data in the form of graph it becomes much easier to visualize and analyses the variation of government responses over a period.

The network formed has sixteen nodes – Eleven for the identified countries across the world and five significant policy indicators (Testing policy, Contact Tracing, Restrictions on gathering, International Travel and Vaccination policy). All the acknowledged countries have attributes values that represents the value of stringency index and containment health in the time series way.

Network has maximum of 55 edges that connects all the countries with policy indicators respectively and edge weight represents the value of policy indices employed by its connected country.

TABLE V. GOVERNMENT RESPONSE NETWORK

Category	Co authorship Network
Dataset	https://covidtracker.bsg.ox.ac.uk/
Nodes	Countries, Policies
Edge type	Undirected, Weighted
Network Format	Undirected
Edge weights	Policy Indices value
Number of nodes	11(countries),5(policies)
Number of edges	55 edges
Average degree	6.6221 edges / vertex
Nodes Attribute	Stringency Index, Containment and health Index

1) Layout selection for the time series graph

To select a proper layout for graph is a significant part as it embodies the graph in most presentable and logical format that further effects the investigation of outcomes. Using the Gephi toolkit [25,29], the network is created and is analysed using an inbuild Force Atlas2 algorithm [30]. Force Atlas2 is a force directed approach that follows the principles of repulsion, attraction, and gravity to deliver a high gradation of accuracy for small to large datasets. It is an advanced version of Force Atlas which is very slow in comparison. Force Atlas2 works on the principle of attraction and repulsion where higher level of attraction will compose the nodes and lower attracting nodes are placed at a distance. This force-based approach, has multiple features like gravity, scaling and overlap prevention apart from these, it gives a layout that is more clustered and visually more vibrant. It's advanced value of gravity assistances in towing the nodes towards the centre of the graph and subordinate value supports in dispersion of the nodes. The *scaling feature* sets a revulsion level for the diffusion of graph to make it better decipherable. The *prevent overlap feature* is used avoid overlying among the nodes. This algorithm is finest suitable to deliver a improved outline of graph for visualization and analysis. "Fig. 2" represents the parameters settings used for Force Atlas2 algorithm. "Fig. 3" represents the parameters settings used for another inbuilt layout algorithms, Frutchterman-Reingold algorithm [31] used to provide better visualization to network.

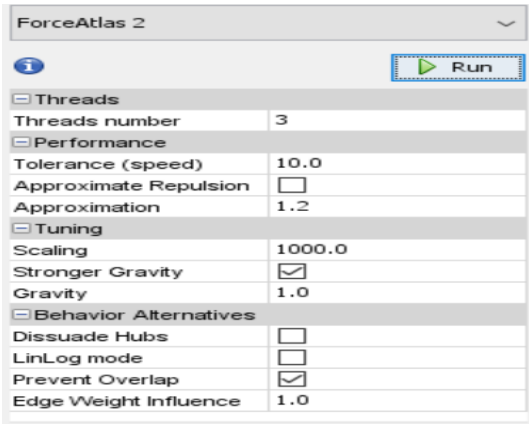


Figure 2. Parameter Settings for Force Atlas2 Algorithm

This algorithm provides a stationary assessment of the graph and operates on the similar principal as Force Atlas 2, the significant functional difference lies underneath, the altering of graph area rather than changing attraction and repulsion parameters. The combination of these two algorithms stretches an improved outline to the graph that can be easily understandable.

"Fig. 4" represents the resulting outline of graph, different colours are given to nodes to discriminate among policy indicators and corresponding nations. (For example, light yellow represents for policy indicators and brown embodies for corresponding nations). Values on edges represents the indices value of each policy implied by its connected nation.

E. Visualization and Analysis of Dynamic Government response network:

1) Visualization of dynamic network: The dynamic network created can be visualized to show the effect of variation in policies indices value on government response index values. These deviations can be easily viewed by applying the subsequent features like variation in size and colour of the nodes and edges based on their attribute values.

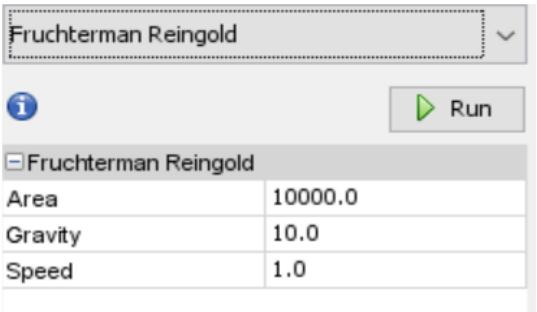


Figure 3. Parameter Settings for Furtcherman-Reingold Algorithm

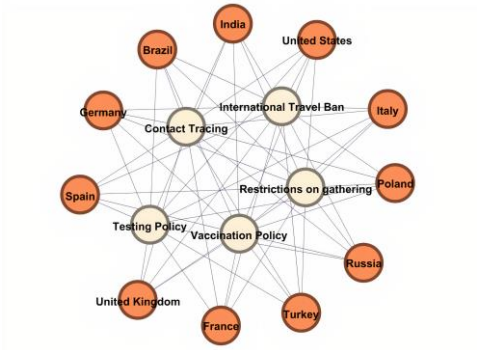


Figure 4. Overview of generated Graph with all nodes and edges.

a) Change of node size/color based on attribute value: The size and colour of a node determines the attribute (response index) [2,32] values in time variant fashion. Through the assistance of dynamic graph, it becomes best possible to analyse the size and colour intensity variations dynamically with respect to the variation in the pragmatic attribute value.For example "Fig. 5" illustrates these variations in nodes with respect to the stringency government policy index. The feature helps to discover the evolution of stringency and containment health index value change over time.

b) Change of edge color and width bottomed on attribute value: The edge colour and width can similarly be altered grounded on its attribute value. As the value of the edges are - indices value of each policy, this feature helps to novel the variation of indices value over time. "Fig. 6" shows how edge colour and width changes with variation in the policy indices.

c) Identification of connected nodes: The visualization and analysis tool offer a feature that shows the node sensible features specifically, its attribute value, count of connected nodes and the edge attribute value [25,33]. This feature further supports in the country wise analysis or policy wise analysis at any point of time.

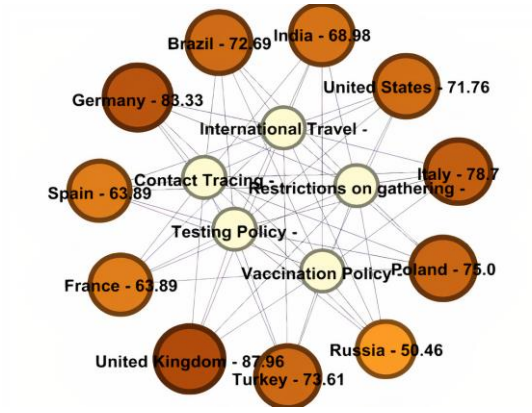


Figure 5. Variation in node size and colour grounded on attribute - stringency government policy index.

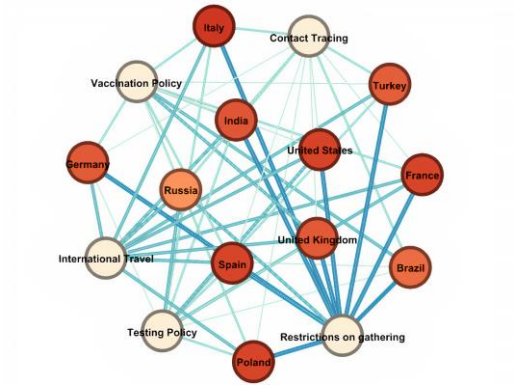


Figure 6. Variation in edge colour width based on weight (policy indices)

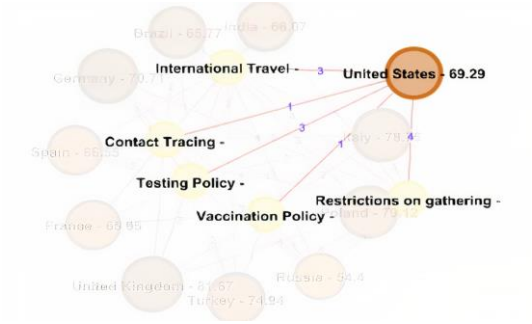


Figure 7. United State with containment and health index value(attribute) and indices value(weight) for observed policies.

*"*Fig. 7*"* shows the country wise visualization respectively It demonstartes the policies implemented in The United States in Dec 2020,in response to Covid-19 and the values on edges specify the level of implementation like level of restriction on gathering was 4 whereas testing policy was implemented on level 3,contact tracing was on level 1,international travel ban was on level 3 and vaccination policy was in its initial phase of level1.Similarly graph can be viewed on the policy basis where all countries implementing that policy can be viewed with the level of implementation intensity.

*2) Analysis of Dynamic Government response network:*To analysis the dynamic graphs various methods and metrics are essential to address the temporal feature of these networks. The most common technique to cater with the temporal aspect is to simply divide the entire network over time into a series of static subgraphs by discretization of dynamic graph [26]. For example for a given graph G in a time period from t(1) to t(m-1), m different subgraphs such as G1,G2.---Gm represents the state of graph at timestamp t1,t2,…..up to tm-1.G1 represents the state of the graph in time interval t1 to t2.Similarly,G2 represents the state of network in time interval t2 to t3 and so on. The term time-stamp is used to refer to individual state of graph(subgraph) according to time periods sampled from G. Such discretization of the dynamic graph into a series of static subgraphs can enable the implementation of many statistical analysis metrics to be applicable on dynamic networks [35,36]. In the present research, statistical analysis is performed to analyse nodes and edges based on their attributes and weights respectively and th find the modularity and weighted degree of the network and explains the importance of this analysis for the better understanding of the process and retrieval of innovative information from the network.

a) Node based Analysis: A graph(G) is a pair of (V, E) where V is the set of vertices and E is the set of edges that connects the two distinct nodes(u,v).A dynamic graph can be node dynamic graph or a edge dynamic graph or in special cases, can be customized to perform both functions. A Node dynamic graph is a graph where the set of vertices V varies with the time and if these vertices(nodes) have some attribute value then it will dynamically be varied over time. The analysis of this dynamics can provide a useful information from a complex network. As in visualization section it can be seen that how nodes colour and size vary with respect to the attribute value over time. In the present research, the nodes specify the countries and government policy indicators. The attribute values of country nodes represent the administration response index i.e. stringency and containment and health towards the pandemic spread across the world.The index value at any point of time depends on the value of policy indices, denoted by edges between countries and policy indicators respectively.Now, the stimulating part here is to analyse the variation in government response index based on the policy indices. This variation can be productively analysed using the dynamic graphs.

The variation is an outcome of varying attribute values (stringency or health index) on the observed nodes. Substantial size and intense colour depict that the attribute value on the detected node is larger as compared to other nodes in the network.Relationship can now be established as - Higher the attribute value enormous is the size and more intense is the colour of the node.Since attributes are representing government response index, it directs to clearly identify the country that has higher response index at a given time and the countries that has the least response indexes. "Fig. 8" illustrates the scenario of May 2020 for the government stringency index across the world. India, an Asian country had the highest stringency index of 96.3, preceding the Italy with second highest stringency index index value of 93.5 and United States with the least stringency index of 72.69.Similar analysis can also be

e

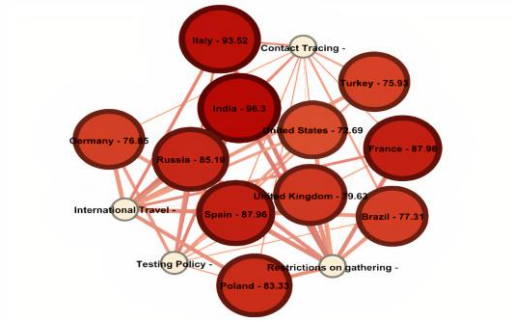


Figure 8. Stringency index in May 2020

done on another attribute i.e health index of countries. "Fig. 9" shows the situation in NOV 2020 for the health and containment index. It can be clearly identified that France had the highest containment and health index of approximately 76 as compared to other countries. Italy had second highest index of 69.94 and Russia had the least index of about 53. This analysis helps to find answers for the identified objectives RQ1 and RQ2 where the variation in node(country) size and colour resembles the effect of policies on government response index. The result obtained from the extensive analysis are discussed in Section 4 for the future reference. The deduced information after the analysis is decisive and significant for the forthcoming planning to face the similar type of situations and dynamic graph makes it more vibrant.

b) Edge based Analysis: As discussed in previous section, an 'edge' is a connection between the two distinct nodes (u, v). Edge can be *weighted* or *unweighted*. A Weighted edge has an additional value (weight) associated with it that specifies the strength or significance of connection between the nodes. Dynamic graph can be *edge dynamic* where set of edges and their associated weight varies with respect to time. In this study, edge represents the connection between countries to be observed and the government policies formed during the pandemic across the globe.

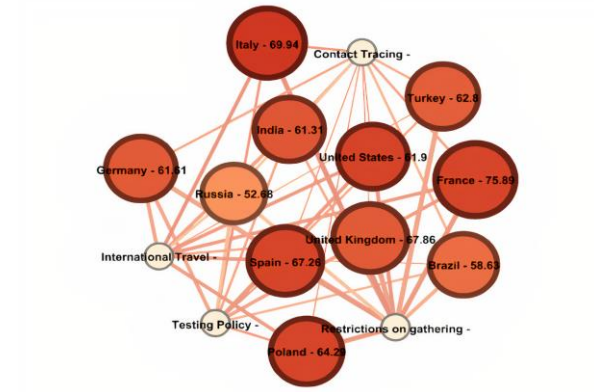


Figure 9. Health Index in Nov 2020

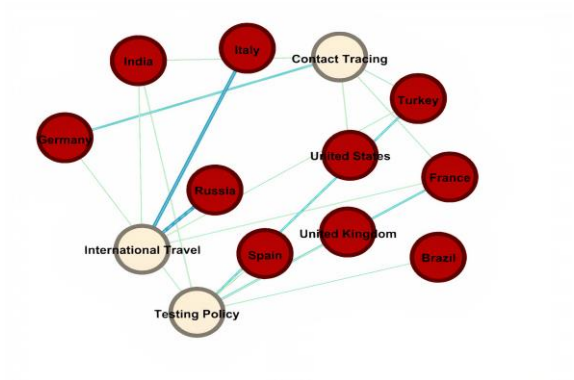


Figure 10. Policies implementation in Feb 2020

The weight of the edges represents the indices value of the policies, in other terms the intensity or severity to which these policies are followed. Higher the value of indices, more sternly policies are being followed. As conversed previously, the edge colour and width may be changed depending on the weight with reference to the varying time. The feature plays a significant role in analysing the policies and their implementation in the republics across the globe. For better visualization, the ranking process is restricted to execute on the set of edges only and nodes are exclusive.

"Fig. 10" shows the scenario in February 2020, where all the policies were in preliminary phase and policies that starts in early phases of pandemic like restriction ongathering and vaccination was not even started at that time. International travel ban was strictly implemented in Italy and Russia, the testing on the other hand was well instigated in United states, Spain, UK, Turkey and France. At that time, countries like India, Germany and Italy were in initial phase of these policies. In contrast, "Fig. 11" shows the situation of February 2021 where the government policies like restriction on gathering is well implemented in all countries.

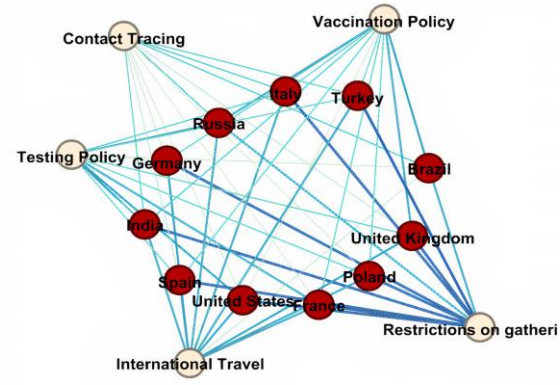


Figure 11. Policies implementation in Fen 2021

The vaccination policy got started and international travel ban becomes less intense. This analysis of transition in one-year aids to establish the replies for objectives RQ1 and RQ2 and also helps in getting more insight for RQ3 where wider and intense edges lead to higher government response index and RQ4 to find significance of each policy especially vaccination.

c) Modularity Class: Dynamic graph supports cluster representation to find the modularity of the graph [35]. Modularity assistances in spotting the existing groups in the graph. A group is a set of nodes that are coupled with ample strength as equated to the further nodes of the graph. Since, the graphs under research are time series graphs, it turns out to be interesting that how number of clusters behave with respect to time. The Blondel's modularity algorithm [37] is simulated to discover modularity classes with a randomise approach and with a resolution of 1.0(a feature to avoid weak groups, higher the value of resolution lesser is the number of groups). Different communities are detected dynamically in the graph based on the edge weight [27] that illustrates significant policies for some nations as likened to others. "Fig. 12" shows the node count distribution and detected two communities of size 7 and 8 nodes respectively. Identified two communities are sensed in December 2020 and specifies that policies like Imparting vaccination, restricting international travel ban and testing for the virus are more significant in nations like-United states, France, Russia, Spain and Poland while, contact tracing and restraint on assemblies are more momentous in other countries. Modularity score is 0.06 that implies that connection outside communities are not sparse. Modularity analysis helps to meet objective RQ4 by helps us in finding significance of vaccination policy. In result section it can be seen that vaccination policy is significant for most of the countries and its implementation leads to improved government response index.

d) Weighted degree: The weighted degree of a node considers the sum of weights of all the connected edges to a node- directed or undirected [32,35,38]. In our research weights on edges represents the policy indices value, that interestingly leads to find weighted degree of each policy and to analyse the severity in the implementation of observed policy at any point of time. "Fig. 13" shows the average weighted degree node distribution of weighted degree and the variation in size of policy nodes based on their weighted degree value in March 2021. This variation clearly indicates that testing policy is least implemented in all policies and international travel ban is the policy that is best implemented. (Average Weighted Degree: 11.875). This analysis of weighted degree helps to meet the objective RQ3 to find the significance of each policy, how well the policy is implemented? and where the improvement is

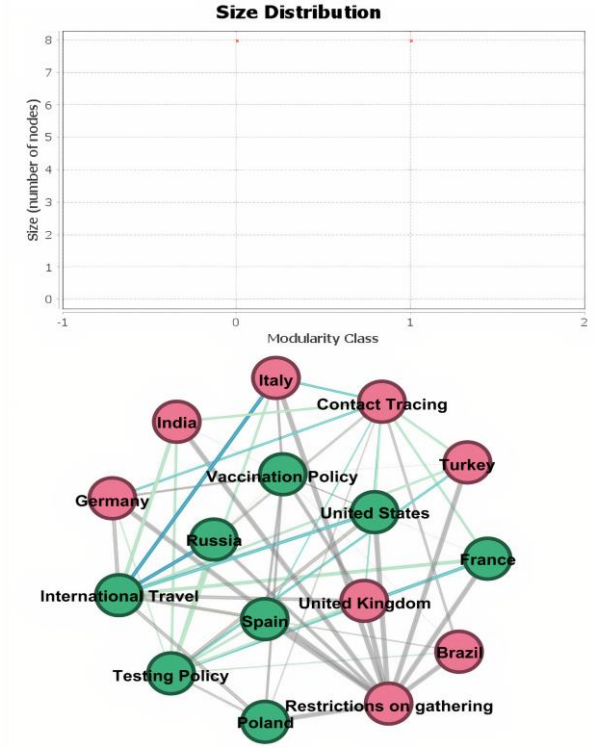


Figure 12. Node count sensible chart for Modularity class and detected communities in Dec 2020.

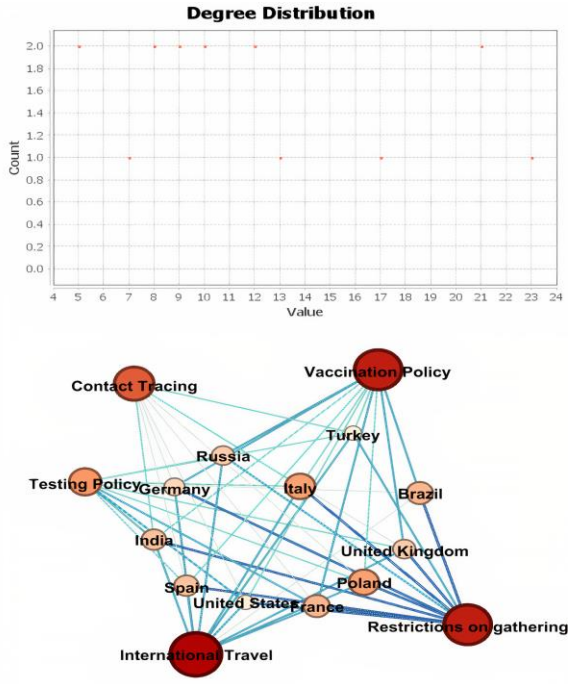


Figure 13. Weighted degree distribution and variation in policy nodes based on weighted degree in March 2021.

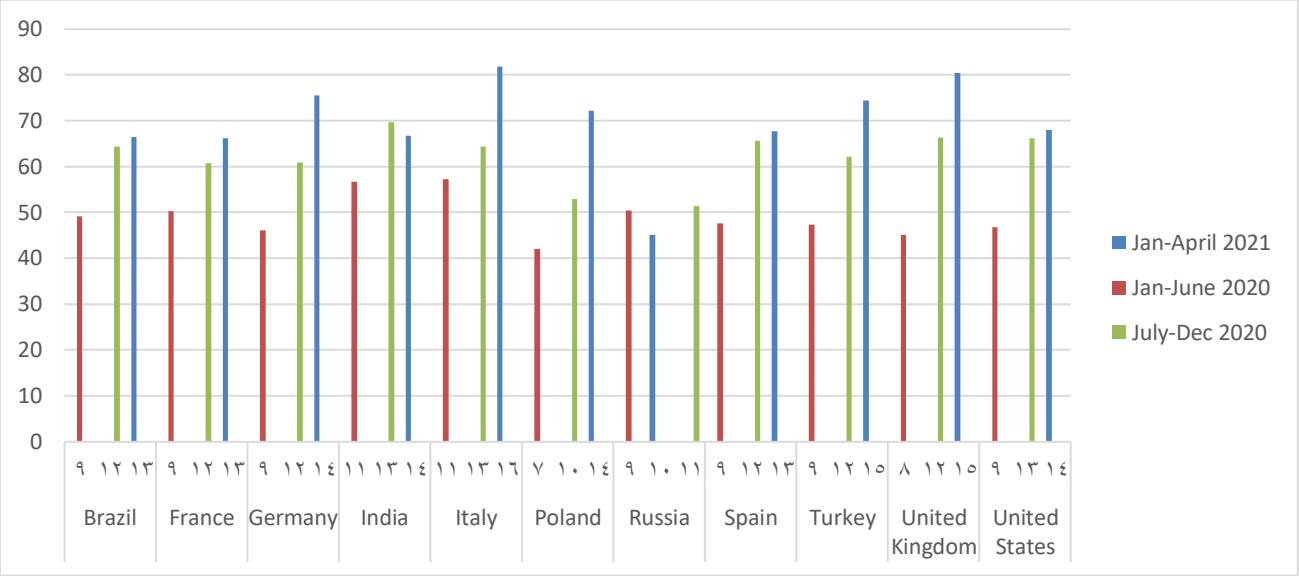


Figure 14. Average of government responses based on average of policy indicators.

required? Improved implementation of these policies can lead to the prevention of complete lock down in similar situations of pandemic. Similar Analysis can be done for any point of time to get better insight of the implementation depth of each policy and its effect on the government response indexes of each country

Based on the above analysis, we tried to show the variation in government responses based on policies. Figure 14 shows the average of both government responses: Stringency and Containment health concerning the par of all considered policies: Restriction, Contact tracing, Testing, International travel ban, and Vaccination. It can see that the higher the value of policy indicator better is the government response toward pandemic.

4. RESULTS AND DISCUSSIONS

Since the color and size of nodes varies conferring to two dimensions of pragmatic response index values namely, stringency index and another is health and containment health. By analysing the designed dynamic graphs in section 3, following results can easily be interpreted.

A Variation in Stringency Index with respect to policies implementation: As per the Oxford University definition *"The stringency index records the strictness of government policies. It does not measure or imply the appropriateness or effectiveness of a country's response*" [20]. However, studies through various paradigms demonstrates that nations with severer administration retorts have experienced gentler epidemic progress tolls in comparison with the countries which are less stringent "Fig. 15" and "Fig. 16" illustrates how the government stringency index varies in accordance to the policy index values with respect to the variation in time domain.

B. Variation in Containment health Index with respect to policies implementation: Containment is commenced in the initial phases of the outburst of novel corona virus and purposes to touch and separate all infested persons as a measure to break the virus from spreading. However, after a short period of time it becomes impossible to comprehend the virus, then exertions passage towards the mitigation stage. At thisstage, actions are engaged to slow down the blowout of virus and mitigate its aftermaths on the healthcare structure and humanity. An integration of both containment and mitigation measures may be commenced at the same time across the globe to improve the situation towards the pandemic.

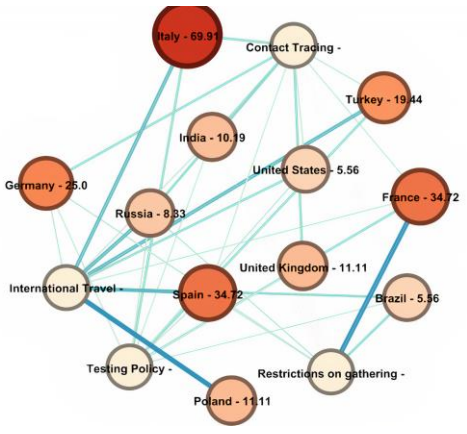


Figure 15. Strnngency Index in March 2020 based on policy indices.

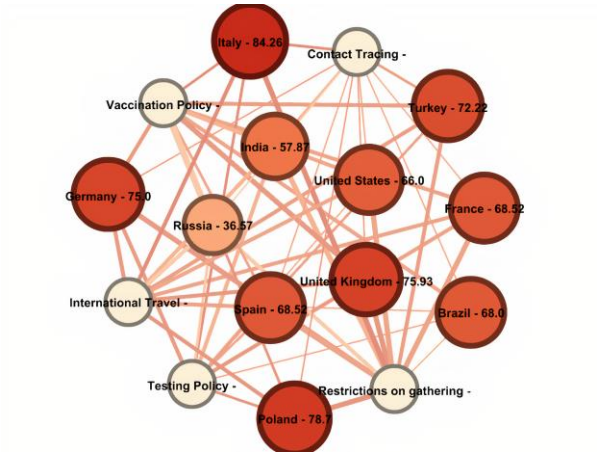


Figure 16. Stringency Index in March 2021 based on policy indices.

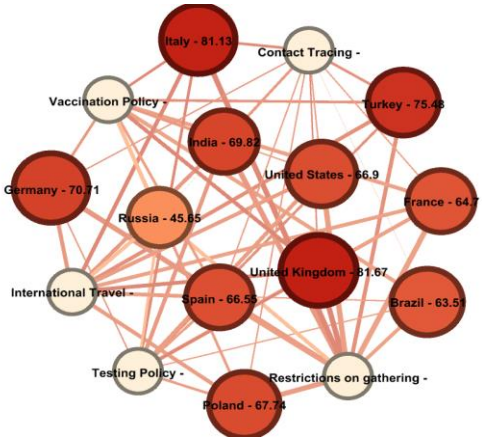


Figure 18. Affluence of Containment health in March 2021

"Fig. 17" and "Fig. 18" shows the variation in containment and health index at different timestamps.As the policy indices varies with time it leads to the variation in helath indexes which clearly visualize as variation in size and colour of country nodes.

C. Implication of each policy and outcome of their enactment on the anticipation of complete lockdown: The indices values considered as policy indicators in the current research are very crucial and are directly impacting the prevention of complete lockdown in a nation that will further influence the economic welfare of any country.Also, an indicator viz. vaccination policy of the nation is examined to demonstrate the outcome of realizing vaccination drives in each country that added primes to plentiful sooner results in improving index values.As previously discussed the impact of significant policies and their implementation depth on indexes as in "Fig. 19" which shows that how both stringency index and containment health index increase in most of the countries as the implementation of all the policies improves.

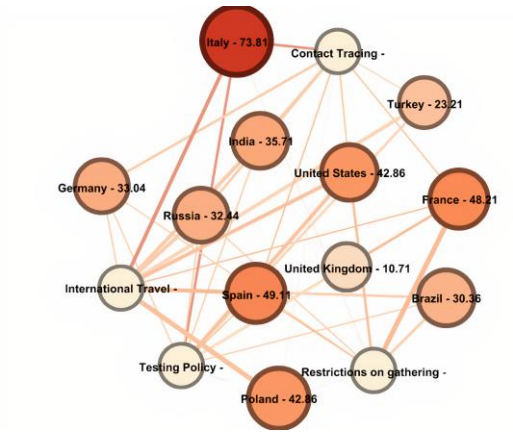


Figure 17. Affluence of Containment health in March 2020

Based on these results it can be stated that if these policies will be implemented by the authorities with more severity then the situation/requirement of complete lockdown can be prevented that also results in much less economic loss to any country.

D. Effect of vaccination policy in improving government responses to Covid-19: It's a well-known fact that there is a vaccine shortage across the world specially in poorer countries. The developed and middle-income countries are doing much better in the inoculation campaigns against the novel corona virus. Till April 2021, 0.93 billion doses have been administrated against Covid, only 2.79% people around the globe are fully vaccinated. Although more than a dozen of vaccines has been created to fight against this pandemic, distributing them equally and to everyone on the planet seems to have become the biggest challenge. Now WHO has developed a framework for production and distribution of vaccines, but the developed countries and major vaccine manufacturers are against joining this initiative.If the current pandemic situation is analysed closely, it is found that many vaccines are developed by public funding agencies. For example, the most popular vaccine AZD_{1222} was developed at oxford university and was intended to use freely by all but later university gives the vaccine marketing rights to British company Astrazeneca. Astrazeneca further signed a contract with Serum institute of India Private Limited to manufacture and market vaccine as Covishield. Wealthy countries that can afford the vaccine have booked more than 2/3rd of vaccine doses that will be produced globally by the end of 2021. Remaining 1/3rd of the vaccine must be share among 92 impoverished nations across the world.Presently, China, USA, India, EU and UK are major manufacturers. The US and UK are preserving medicines for domestic use only while EU only shares vaccines among its member states. India and China are two major countries that are

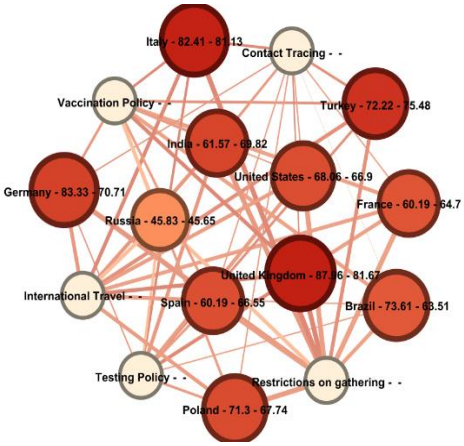


Figure 19. Improvement in both stringency and health indexes in Feb 2021

exporting vaccines to developing world. India is providing the vaccine on the roadmap of "Vaccine diplomacy" to 70 countries across the globe. This helps the country to strengthen domestic ties with many nations, still it is facing huge domestic shortage of vaccines.Since it is implied that there is a major role of vaccination policy in improving the response indexes of any country. It is necessary that governments across the globe should necessary nourish the manufacturing institutes to ease the requirement.

"Fig. 20" shows that after the implementation of vaccination policy in Jan 2021 the index value of both stringency and containment and health increases after the introduction of vaccination policy in Dec 2020 in many countries.Although this indicator is not used in calculating the response index but by analysing its effect in graph it can be concluded that this policy will play a major role in near future for the improvement of the government response indexes in the similar pandemic situation. Similar results can be deduced based on modularity class distribution of communities where nodes of same community are representing in same colour.

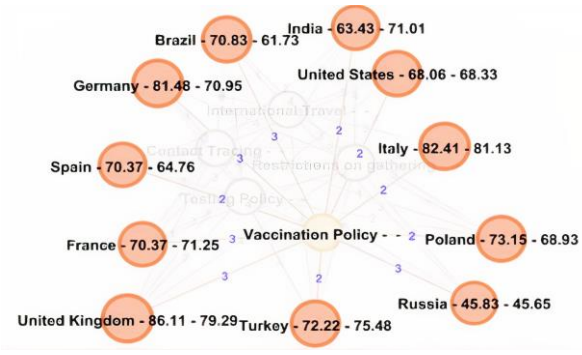


Figure 20. Vaccination policy and its indices value for each country

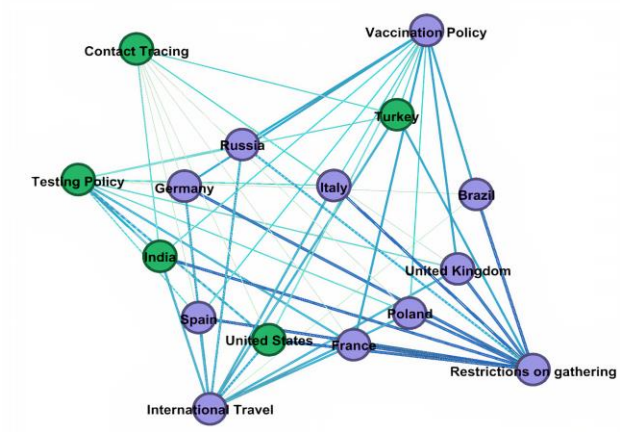


Figure 21. Modularity class distribution of communities in Feb 2021

"Fig. 21" shows the result of modularity analysis in Feb 2021 and demonstrates how significant was vaccination policy for most of the nations it can be proved the most significant weapon in future to won the battle against current pandemic.

5. CONCLUSION AND FUTURE SCOPE

The illustrative application is designed to inspire exhaustive exploration and to ease the decision support system for policymakers and researchers that further helps them to answer significant communal policies and epidemiological queries on current pandemic situation.OxCGRT dataset is used for logical analysis of the outcomes of administrations' retorts to COVID-19 across diverse populaces and contexts, in conjunction with the description of major factors that influence the governments to adopt different policies.The research further argues that any analysis of an explicit nation should be done on the foundation of the fundamental policies, not on the indices alone. The paper concludes that governments across the world with more stringent policies to respond to Covid-19 pandemic may perhaps give rise to the improved performance that further contributes to curve the death tolls.Similar study can be implemented on states or district level to get better insight of effect of novel corona virus or any pandemic fighting policies in interior portions of the country that are otherwise unreachable for analysis.

COVID-19 is still at the gauge of pandemic, implies that authorities and government sectors altogether are experiencing the same 'exigency' stage. In the future, we intend to focus on the behavioural aspect of government policies to handle the pandemic situation. For example, the use of appropriate language used in communication by the government personals, to enhance the citizens' morale and help them recover from psychological distress from the Covid-19 pandemic. The government policies should be

designed that try to upsurge the self-esteem, essence to contest, and harmony among the people by bringing encouraging expressions, conveying optimistic posts and alluring the citizens to support each other. We intend to analyse the patterns among the speeches of various diplomats across the world and will try to establish the relationship between the citizens positive morale in accordance with the linguistic behaviour of the authorities.

REFERENCES

[1] Ubaru, S., Horesh, L., & Cohen, G. (2020). Dynamic graph based epidemiological model for COVID-19 contact tracing data analysis and optimal testing prescription. arXiv preprint arXiv:2009.04971.

[2] Aggrawal, N., & Arora, A. (2016, October). Visualization, analysis and structural pattern infusion of DBLP co-authorship network using Gephi. In 2016 2nd International Conference on Next Generation Computing Technologies (NGCT) (pp. 494-500). IEEE

[3] Gupta K.D., Dwivedi R. and Sharma D., "Prediction of Covid-19 Trends in Europe using Generalized Regression Neural Network Optimized by Flower Pollination Algorithm," Journal of Interdisciplinary Mathematics, DOI: 10.1080/09720502.2020.1833447.

[4] Banyal S., Dwivedi R. , Gupta K.D. , Sharma D., Al-Turjman. F and Mostarda L., "Technology Landscape for Epidemiological Prediction and Diagnosis of COVID-19,"CMC-Computers, Materials & Continua, Vol.67, No.2, pp. 1679-1696, 2021.

[5] Mojjada, R. K., Yadav, A., Prabhu, A. V., & Natarajan, Y. (2020). Machine Learning Models for covid-19 future forecasting. Materials Today: Proceedings.

[6] Wang, P., Lu, J. A., Jin, Y., Zhu, M., Wang, L., & Chen, S. (2020). Statistical and network analysis of 1212 COVID-19 patients in Henan, China. International Journal of Infectious Diseases, 95, 391-398

[7] Das N. N. et al. Automated Deep Transfer Learning-Based Approach for Detection of COVID-19 Infection in Chest X-rays, IRBM Elsevier, ISSN 1959-0318, 2020.

[8] Kumar N. et al. Efficient automated disease diagnosis using machine learning models, in Journal of Healthcare Engineering, ISSN: 2040 – 2295, 2021.

[9] Min C, Shen F, Yu W, Chu Y. The relationship between government trust and preventive behaviors during the COVID-19 pandemic in China: Exploring the roles of knowledge and negative emotion. Prev Med. 2020 Dec;141:106288. doi: 10.1016/j.ypmed.2020.106288. Epub 2020 Oct 19. PMID: 33091414; PMCID: PMC7571476.

[10] Ma C, Ma L, Helwan A, Ma'aitah MKS, Jami SA, Mobarak SA, Das NK, Haque MA. An online survey and review about the awareness, coping style, and exercise behavior during the "COVID-19 pandemic situation" by implementing the cloud-based medical treatment technology system in China among the public. Sci Prog. 2021 Apr-Jun;104(2):368504211000889. doi: 10.1177/00368504211000889. PMID: 33827338.

[11] Sultan, & Rapi, M. (2020). Positive discourse analysis of the indonesian government spokesperson's discursive strategies during the covid-19 pandemic. GEMA Online Journal of Language Studies, 20(4), 251–272. https://doi.org/10.17576/gema-2020-2004-14

[12] Haider, S. H., Asad, M., & Aziz, A. (2015). A survey on the determinants of entrepreneurial training effectiveness among micro finance institutions of Malaysia. Mediterranean Journal of Social Sciences, 6(6 S4), 396-403. https://doi.org/10.5901/mjss.2015.v6n6s4p396

[13] Desalegn Z, Deyessa N, Teka B, Shiferaw W, Hailemariam D, Addissie A, et al. (2021) COVID-19 and the public response: Knowledge, attitude and practice of the public in mitigating the pandemic in Addis Ababa, Ethiopia. PLoS ONE 16(1): e0244780. https://doi.org/10.1371/journal.pone.0244780

[14] Angaw KW. Policy Responses and Social Solidarity Imperatives to Respond the COVID-19 Pandemic Socioeconomic Crises in Ethiopia. Clinicoecon Outcomes Res. 2021 Apr 22;13:279-287. doi: 10.2147/CEOR.S300695. PMID: 33935506; PMCID: PMC8076721.

[15] Hadi, S., & Supardi. (2020). Revitalization strategy for small and medium enterprises after corona virus disease pandemic (Covid-19) in Yogyakarta. Journal of Xi'an University of Architecture & Technology, 8(4), 4068-4076.

[16] Im, U.L., Lam, C.-C.C. and Ma, E. (2021), "The effective responses of a tourism and gaming city towards the devastating effects of pandemic: a case study of Macao", International Journal of Tourism Cities, Vol. 7 No. 2, pp. 492-510. https://doi.org/10.1108/IJTC-10-2020-0219

[17] M. K., Tiwari, A., Chu, A. M., Tsang, J. T., & Chan, J. N. (2020). Visualizing COVID-19 pandemic risk through network connectedness. International Journal of Infectious Diseases, 96, 558-561.

[18] Wajahat, A., Nazir, A., Akhtar, F., Qureshi, S., Razaque, F., & Shakeel, A. (2020, January). Interactively visualize and analyze social network gephi. In 2020 3rd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET) (pp. 1-9). IEEE.

[19] Alguliyev, R., Aliguliyev, R., & Yusifov, F. (2021). Graph modelling for tracking the COVID-19 pandemic spread. Infectious Disease Modelling, 6, 112-122.

[20] Gnanvi, J., Salako, K. V., Kotanmi, B., & Kakaï, R. G. (2021). On the reliability of predictions on Covid-19 dynamics: A systematic and critical review of modelling techniques. Infectious Disease Modelling.

[21] Shuja, J., Alanazi, E., Alasmary, W., & Alashaikh, A. (2020). COVID-19 DATASETS: ASURVEY AND FUTURE CHALLENGES. development, 11, 12.

[22] Thakur, H. K., Gupta, A., Nag, S., & Shrivastava, R. (2020). Multi-class instance-incremental framework for classification in fully dynamic graphs. International Journal of Computational Science and Engineering, 21(1), 69-83

[23] Khan, M. A., & Atangana, A. (2020). Modeling the dynamics of novel coronavirus (2019-nCov) with fractional derivative. Alexandria Engineering Journal, 59(4), 2379-2389.

[24] Azad, S., & Devi, S. (2020). Tracking the spread of COVID-19 in India via social networks in the early phase of the pandemic. Journal of Travel Medicine, 27(8), taaa130

[25] Bokhare, A., & Metkewar, P. S. (2021). Visualization and Interpretation of Gephi and Tableau: A Comparative Study. In Advances in Electrical and Computer Technologies: Select Proceedings of ICAECT 2020 (pp. 11-23). Springer Singapore.

[26] Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. Social networks, 32(3), 245-251.

[27] Hussain, S., Muhammad, L. J., & Yakubu, A. (2018). Mining social media and DBpedia data using Gephi and R. Journal of Applied Computer Science & Mathematics, 12(1), 14-20.

[28] Hale, T., Petherick, A., Phillips, T., & Webster, S. (2020). Variation in government responses to COVID-19. Blavatnik school of government working paper, 31, 2020-11.

[29] Korobkov, I., Korobkov, D., Gurjanov, A., & Shukalov, A. (2021). Comparison of Graph Visualization Tools for the Problem of Evaluating the Characteristics of SpaceFibre Networks. In Conference of Open Innovations Association, FRUCT (No. 28, pp. 578-581). FRUCT Oy.

[30] Jacomy, M., Venturini, T., Heymann, S., & Bastian, M. (2014). ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. PloS one, 9(6), e98679.

[31] Fruchterman, T. M., & Reingold, E. M. (1991). Graph drawing by force - directed placement. Software: Practice and experience, 21(11), 1129-1164.

[32] Groshek, J., de Mees, V., & Eschmann, R. (2020). Modeling influence and community in social media data using the digital methods initiative-twitter capture and analysis toolkit (DMI-TCAT) and Gephi. MethodsX, 7, 101164.

[33] https://covidtracker.bsg.ox.ac.uk/

[34] Zaidi, F., Muelder, C., & Sallaberry, A. (2014). Analysis and visualization of dynamic networks. arXiv preprint arXiv:1409.5034.

[35] Thakur, H. K., & Gupta, A. (2017). Mining Maximal Quasi Regular Patterns in Weighted Dynamic Networks.In mecs-press(Vol No-4, pp. 48-62)

[36] El Mouden, Z. A., Taj, R. M., Jakimi, A., & Hajar, M. (2020). Towards Using Graph Analytics for Tracking Covid-19. Procedia Computer Science, 177, 204-211.

[37] Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. Journal of statistical mechanics: theory and experiment, 2008(10), P10008

[38] Beck, F., Burch, M., Diehl, S., & Weiskopf, D. (2017, January). A taxonomy and survey of dynamic graph visualization. In Computer Graphics Forum (Vol. 36, No. 1, pp. 133-159).



Author 1 Ms. Neelu Chaudhary is working as an Assistant Professor at department of Computer Science and Technology of Manav Rachna University (MRU), Faridabad. She has more than 13 years of teaching experience. She has completed her B. Tech in 2004 and M. Tech in 2006 from Kurukshetra University, Kurukshetra (KUK). Currently she is pursuing her Ph.D. in Computer Science and Engineering from Manav Rachna University, Faridabad. Her area of research is data mining and dynamic graph mining. She has published 3 papers in International Journals and 4 papers in international conferences**.**



Author 2 Dr. Hardeo Kumar Thakur is working as an Associate Professor at department of Computer Science and Technology of Manav Rachna University (MRU), Faridabad. He has more than 12 years of teaching and research experience in leading institutions of India. He has earned his Ph.D. (Computer Engineering) from University of Delhi in 2017 in the field of data mining.

Dr. Thakur has published 14 research papers in international journal of repute, 13 papers in international conferences and 1 book chapter. His current research interests are Data Mining, Dynamic Graph Mining, Machine Learning and Big Data analytics. He is an active referee for many international Journals and Conferences.



Author 3 Dr Rinky Dwivedi has completed her B. Tech in Computer Science and Engineering from Guru Gobind Singh Indraprastha University, Delhi in 2004 and M.E. in Computer Technology and Application from Delhi College of Engineering, Delhi in 2008. She has received her Doctorate in 2016 from Delhi Technological University, New Delhi in the field of Agile Method Engineering. Her area of research includes Software Engineering and Method Engineering, Agile Software and Machine Learning. Dr. Rinky has over 14 years of experience in Academics. She worked as Assistant Professor for 9+ years at Maharaja Surajmal Institute of Technology, Delhi. She has also worked as a full-time research scholar in Delhi Technological University for two years and is currently she is working as Associate Professor in Maharaja Surajmal Institute of Technology, Delhi. She has published more than 20 research papers in reputed Journals and conferences proceedings and has also authored books.

Wireless Personal Communications
https://doi.org/10.1007/s11277-021-08860-y

Edge Detection in Digital Images Using Guided L_0 Smoothen Filter and Fuzzy Logic

Akshi Kumar[1] · **Sahil Raheja**[2]

Accepted: 8 August 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract
Image segmentation is an important process in computer vision. Recently fuzzy logic based edge detection is heavily investigated as by changing the number of rules edge detection can be improved. However, due to large colour variations in the images false edges are detected and even using fuzzy rules they cannot be reduced significantly. These falsely detected edges can be controlled by using smoothen filter while controlling the degree of smoothness. This paper, presents fuzzy logic based edge detection mechanism while using Guided L_0 smoothen filter for the smoothening of image under various degree of smoothens. Simulation results for edge detection is presented for Canny, Sobel, Fuzzy logic based edge detection and finally fuzzy logic edge detection with inclusion of L_0 smoothen filter. The results are compared with classical and modern methods. Simulation is performed on Berkley Segmentation Database (BSD) and USC-SIPI Image Database while considering more than 100 images. The obtained F-measure is as high as 0.848.

Keywords Edge detection · Guided filter · Fuzzy logic · Image processing · Sparsity · L_0 smoothing filter

1 Introduction

Edge detection is a prominent area of interest in various fields of research and engineering. An edge can be characterized as a collection of associated pixels separating two distinct boundary regions [1]. An Edge could be defined as a local concept yet the boundary could be defined as a global concept. An ideal edge is a set of pixels whose intensity changes abruptly like step function. Blurry edges are likewise obtained by the elements such as issues or imperfections occurred during of sampling, optics, and image acquisition. In this way, we can closely observe that an edge possess a ramp-like profile [2]. The slope of the ramp is associated with the measure of blurriness. The length of the ramp is termed as the thickness of the edge. Sharp edges are thin while blurred edges are thick. If intensity is constant than first derivative is zero, while in

✉ Sahil Raheja
sahilrahejaphd@rediffmail.com

[1] Department of Information Technology, Netaji Subhas University of Technology, Delhi, India

[2] Department of Information and Technology, Delhi Technological University, Delhi, India

Published online: 24 August 2021

ramp intensity profile derivative is constant. Additionally in the 2nd derivative, we can easily notice that it is negative along the light side of the edge while positive along the dark side of the edge such as delta functions. Moreover, we can observe it as zero along and outside the ramp [3].

Edge detection in a grey image is based greatest gradient. In discrete images gradient is defined as pixel intensity difference between any two pixels. Considering 3×3 mask as in Fig. 1a, the center pixel position is denoted by (i, j) and other pixels positions are also shown, correspondingly pixels values are shown in Fig. 1b, and possible edge directions are shown in Fig. 1c.

Let us define pixels positions on pixel mapping is

$$z_1 = f(i-1, j-1), \quad z_2 = f(i, j-1), \ldots \quad \text{and} \quad z_9 = f(i+1, j+1) \tag{1}$$

The resultant image is shown in Fig. 1b. The corresponding edge directions are shown in Fig. 1c. The absolute difference in 135° and 45° can be evaluated as

$$D_1 = |z_5 - z_1| + |z_9 - z_5| \quad \text{and} \quad D_3 = |z_5 - z_3| + |z_7 - z_5| \tag{2}$$

Similarly in 0° and 90° the differences are

$$D_2 = |z_5 - z_2| + |z_8 - z_5| \quad \text{and} \quad D_4 = |z_5 - z_4| + |z_6 - z_5| \tag{3}$$

Finally, edge can be calculated as

$$E_d = D_1 + D_2 + D_3 + D_4 \tag{4}$$

The calculated edges depend on the chosen thresholds and pixel differences. Therefore, if pixel difference and thresholds varies the detected edges also changes. If threshold is high only dominent edges will be detected and if threshold is low, many unwanted edges will also be detcehed. In few applications, where only dominent edges need to be detected, unwanted edges need to be suppressed. The supression of unwanted edges can be done using smoothening of images. In this paper, guided L_0 smoothen filter and fuzzy logic based edge detection method is proposed. Image smoothening is done using guided L_0 smoothen filter and for edge detection fuzzy logic is used.

The rest of the paper is organised as follows: Sect. 2 presents a review of various edge detection mechanisms. In Sect. 3, the proposed edge detection mechanism alongwith mathematical formulation is discussed, and in Sect. 4, of the paper proposed edge dtection framework is described. The simulation results under various edge detection mechanism is reported in Sect. 5 of the paper, and the important conclusions are discussed in Sect. 6.

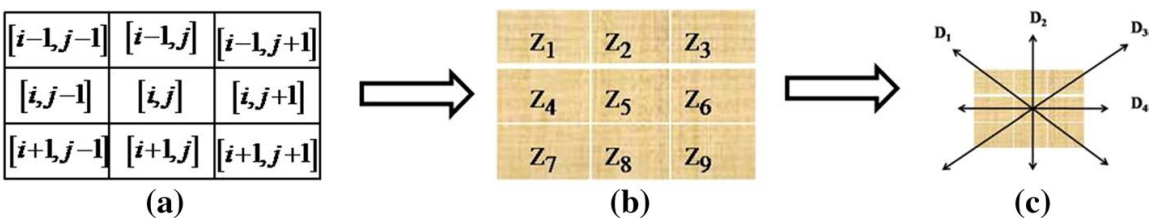


Fig. 1 Traditional edge detection processes **a** 3×3 mask, **b** z notation conversion, **c** possible edge directions

2 Related Works

Edge detection is a prominent area of interest in various fields of research and engineering [2]. Therefore, both edge enhancement and detection is important in various classes of engineering applications. A number of methodologies have been proposed and examined for edge detection. Each one of these was introduced to deal with the limitations of the earlier techniques [3]. The schemes can be characterized as follows:

2.1 Classical Methods

Classical methods are simple and based on the convolution. The image is convolve with a filter or well developed kernels. It is possible to recognize an edge due to the sudden variations in grey scale pixel intensities. The pioneer kernel based methods in edge detection are Robert [4], Canny [5] were proposed. Later on Genming and Bouzong [6] presented a 5×5 kernel based method to locate edges in an image while considering a fixed threshold level. Later on two more kernel based methods Sobel [7], Prewitt [8] were proposed. Later on edge detection method based on Gabor filters were discussed [9, 10]. But results were not promising as fixed threshold was used.

2.2 Classical Learning Algorithms

The classical learning based methods are Mairal et al. [11], gPb (gradient descent) [12], dictionary learning [13], and structured forest [14] (decision trees). These methods outperformed state-of-the-art techniques based on low-level processes in the BSDS segmentation dataset at the time they were proposed [13], achieving the best F-measure values. In most of the cases results were good, but still these methods fail in complex edges.

2.2.1 Soft Computing Algorithms

Ongoing researches fuses techniques based on artificial and swarn intelligence like ant colony optimization (ACO) [15], artificial neural networks [16], particle swarm optimization (PCO), genetic algorithms (GA) [17] and cuckoo search optimization [18]. In recent past, hybrid methods based on ACO and guided image filter were also proposed, in these works edge enhancement is done using guided image sharpening and various forms of ACO were used for the edge detection [19, 20].

2.2.2 Deep Learning Algorithms

Recently, Convolutional Neural Networks based methods have gain popularity in edge detection [10, 21–25]. An architecture based on VGG16 [26] and pre-trained with Image Net dataset is proposed in HED [22]. The network constructs a multi-scale learning architecture by generating edges from each convolutional block. For each predicted edge-map, the training method employs a modified cross entropy loss function. Xie and Tu [22] and Liu et al. [24] have offered modifications based on the same design as their backbone. While each output is fed from each convolution from each block in [24], and in [22], a series of fusion backward processes is conducted using the data of each output. In general, most contemporary DL-based

mode lsuse the convolutional blocks of VGG16 architecture as their backbone. Other DNN based methods are Deep-Contour [27], DeepEdge [28], and CSCNN [29]. Holistically method have automatic learning capability are based on deep learning phenomenon.

Other recent methods are Anisotropic Gaussian Kernels [30], Fuzzy cellular automata [31], single pixel imaging was proposed for edge detection in [32].

Recently, it was shown that in the presence of noise/pubturbation CNN based edge detection struggles to find correct edges [33] and accuracy of edge detection reduces significantly. Therefore, learning based methods are also not very effective in presence of noise.

2.2.3 Fuzzy Logic Based Approaches

In past, concept of fuzzy reasoing for the dege detection was discussed in [34–40]. Kim et al. discussed a method based on fuzzy rules; here 16 fuzzy rules were devised to characterize edges [37]. The performance of the edge detection method is quite accurate for noise free images, but in the presence of noise falsely detected edges rises sharply. More experiments have been performed in higher type of fuzzy logic particularly fuzzy type-2 to oblige more noteworthy vulnerabilities [41, 42]. The structure of type-2 fuzzy logic is complex and not much improvement is observed in edge detection.

In our recent work, we have shown that sharpening of image using guided filtering can improve edge detection significantly [43]. However, to correctly detect edges, sometime false edges need to be suppressed. This can be achieved using Guided L_0 Smoothen Filter. The method proposed in this paper cosnider type-1 fuzzy logic for edge detection, and guided L_0 image filtering for the smoothening of the images. The main of smoothening is to reduce the strength of the undesired edges. The smoothening also very effective in case of small perturbation where deep learning based methods fails very easily. The fuzzy logic is very useful in edge detection because it decision making capability between partial true and partial false with respect to true and false edges.

3 Proposed Method

Smoothness of an image can change detected edges, as smoothness reduces amplitude variation therefore unwanted edges can be diminished. In the spirit of edge preserve smoothening, various methods are proposed over the period of time. In the similar context L_0 smoothing filter is proposed, in this method prominent edges are preserved by increasing the steepness of transition and diminishing the other edges, still maintaining the overall structure of an image. After smoothness of the image fuzzy logic based edge detection method is applied. For better understanding of the Guided L_0 smoothen filter we first discuss L_0 smoothing filter, thereafter L_0 gradiant minimization is detailed which sharpen the dominant edges. Finally, Guided L_0 smoothen filter is discussed which has the edge-preserving smoothing property. Therefore, Guided L_0 smoothen filter takes the advantages of L_0 gradiant minimization and guided filtering.

3.1 L_0 Smoothing Filter

In this subsection, basic concept of L_0 smoothing filter as discussed in [44]. In this method L_0 norm which define non-zero difference in pixel values, and the count (N) of non-zero difference i.e., $\left|z_{i,j} - z_{i-1,j}\right| \neq 0$ in pixel values can be written as

$$c(I_i) = N\{\, i,j|\ |z_{i,j} - z_{i-1,j}| \neq 0\} \tag{5}$$

where i, j and $(i-1), j$ represents neighboring samples (or pixels) indices. $|z_{i,j} - z_{i-1,j}|$ is the forward difference of the intensity also known as gradient w.r.t. i. This is, the L_0 norm of gradient. We consider objective function as (omitting index j)

$$\min_I \sum_i (I_i - I_s)^2 \text{ s.t.,} \quad c(I_i) = \alpha \tag{6}$$

where $c(I_i) = \alpha$, denotes the number of non-zero gradient and the input discrete signal is denoted by variable I_i and its smoothed version is represented by variable I_s. Thus, more appropriate objective function which represents the constraint optimization would be

$$\min_I \sum_i (I_s - I_i)^2 + \lambda\, c(I_i) \tag{7}$$

This optimization is necessary to maintain the image structure as the value of α may be very large. The parameter λ is very important as it controls the sparsity of the image gradient and consequently the smoothness of the output image.

3.2 L_0 Gradient Minimization

The L_0 norm might be directly optimized for a piecewise stable output picture with the aid of L_0 gradient minimization [44]. It sharpens noticeable edges by increasing transition steepness. As a result, problem (8) must be solved after minimization.

$$\min_I \|I - I^*\|_2^2 + \lambda \|\nabla I\|_0 \tag{8}$$

Referring Eq. (8), output smoothened image is denoted by I, and its gradients as ∇I the parameter I^* is the observed image, and λ is a weight to control gradient. With the end purpose to overcome the issue of the objective function and to deal the discontinuity of the term $\|\nabla I\|_0$, an auxiliary variable Ω is introduced, therefore (8) can be written as:

$$\min_I \left\{ \|I - I^*\|_2^2 + \beta \|\Omega - \nabla I\|_0 \right\} + \lambda \|\Omega\|_0 \tag{9}$$

The similarity between ΔI and Ω is controlled by parameter β and the degree of smoothness is handelled by λ.

3.3 Guided L_0 Smoothing Filter

In Fig. 2, detailed description of guided L_0 smoohen filter is shown. In this process, first image is passed through guided filter thereafter L_0 filter, and known as guided L_0 filtering and when min operation is applied on input image and guided L_0 filtering, the resultant image is defined as guided smoothen image.

Xu et al. introduced L_0 gradient minimization, which sharpens the image while keeping dominating edges [44]. Ding et al. propose a guided L_0 smoothing filter. It makes advantage of both the L_0 gradient minimization and guided filter capabilities [45].

To begin, Using, I^k, the parameter Ω^k is optimized using

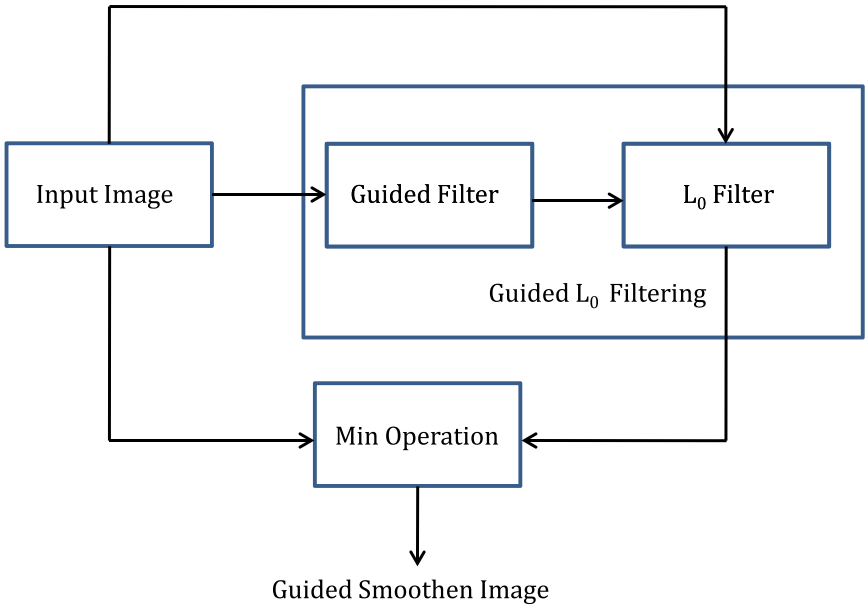


Fig. 2 Guided L_0 smoothen filter

$$\min_{\delta^k} \beta^k \left\| \nabla I - \Omega^k \right\|_2^2 + \lambda \left\| \Omega^k \right\|_0 (k = 1, 2, 3....) \tag{10}$$

As detailed in [45], the parameter, Ωk can be defined as

$$\Omega^k = \begin{cases} 0 & \nabla I^k \leq \frac{\lambda}{\beta} \\ \nabla I^k & others \end{cases} \tag{11}$$

Now both Ω^k and I^k are known, we now evaluates I^{k+1} using Eqs. (9) and (11):

$$\min_{I^{k+1}} \left\| I^{k+1} - I^* \right\|_2^2 + \beta^k \left\| \Omega^k - \nabla I^{k+1} \right\|_2^2 \tag{12}$$

The expression (12) while considering [Eq. (11)] can be written as

$$\begin{cases} \min\limits_{I^{k+1}} \left\| I^{k+1} - I^* \right\|_2^2 + \beta^k \left\| \nabla I^{k+1} - H * \nabla I^k \right\|_2^2 \\ H = \begin{cases} 0 & \Omega^k = 0 \\ 1 & \Omega^k \neq 0 \end{cases} \qquad (k = 1, 2, 3, ...) \end{cases} \tag{13}$$

The solution of Eq. (13) using least squre and fast Fourier transform (F_r) as detailed in [33] is

$$I^{k+1} = F_r^{-1} \left(\frac{F_r(I^*) + \beta \left(F_r\left(\partial_x^T\right) F_r\left(\Omega_x^k\right) + F_r\left(\partial_y^T\right) F_r\left(\Omega_y^k\right) \right)}{F_r(1) + \beta \left(F_r\left(\partial_x^T\right) F_r\left(\partial_x\right) + F_r\left(\partial_y^T\right) F_r\left(\partial_y\right) \right)} \right) \tag{14}$$

In above, inverse fast Fourier transform is denoted by F_r^{-1}, in the horizontal and vertical directions, ∂_x and ∂_y are difference operators.

Defining new variable 's' for smoothed image, then using, Ω^k and s^k we obtain s^{k+1} as

$$\begin{cases} \min\limits_{s^{k+1}} \left\| s^{k+1} - s^* \right\|_2^2 + \beta^k \left\| \nabla s^{k+1} - H * \nabla s^k \right\|_2^2 \\ H = \begin{cases} 0 & \Omega^k = 0 \\ 1 & \Omega^k \neq 0 \end{cases} \qquad (k = 1, 2, 3...) \end{cases} \tag{15}$$

The solution of (15) is:

$$s^{k+1} = F_r^{-1}\left(\frac{F_r(s^*) + \beta\left(F_r\left(\partial_x^T\right) F_r\left(H. * \nabla s_x^k\right) + F_r\left(\partial_y^T\right) F_r\left(H. * \nabla s_y^k\right)\right)}{F_r(1) + \beta\left(F_r\left(\partial_x^T\right) F_r\left(\partial_x\right) + F_r\left(\partial_y^T\right) F_r\left(\partial_y\right)\right)} \right) \tag{16}$$

Algorithm:
Input: Image s^*, guided image I^*, parameters λ, β_0, $\beta_{\max}$, rate κ.
Initialization: $I^1 \leftarrow I^*, s^1 \leftarrow s^*, \beta^1 \leftarrow \beta_0, k \leftarrow 1$,
repeat:
with I^k, solve Ω^k for in (13);
with I^k and Ω^k, solve for I^{k+1} in (14);
with s^k and Ω^k, solve for s^{k+1} in (16);
$\beta \leftarrow \kappa\beta, k++$;
Until $\beta > \beta_{\max}$.
Output: s.

4 Proposed Framework

The basic architecture configuration of the proposed fuzzy expert system for edge detection is shown in Fig. 3. M_1, M_2 and M_3 are the three approaches discussed in this paper. On the input image, only fuzzy logic is applied directly to detect edges in method M_1, and defined as fuzzy logic based edge detection. The input image is first smoothed with the L_0 smoothen filter in method M_2, and then fuzzy logic is employed to find edges and defined as L_0 smoothen filter + fuzzy logic. The input image is first smoothed with the guided L_0 smoothen filter in method M_3, and then fuzzy logic is used to detect edges defined as Guided L_0 smoothen filter + fuzzy logic. It is important

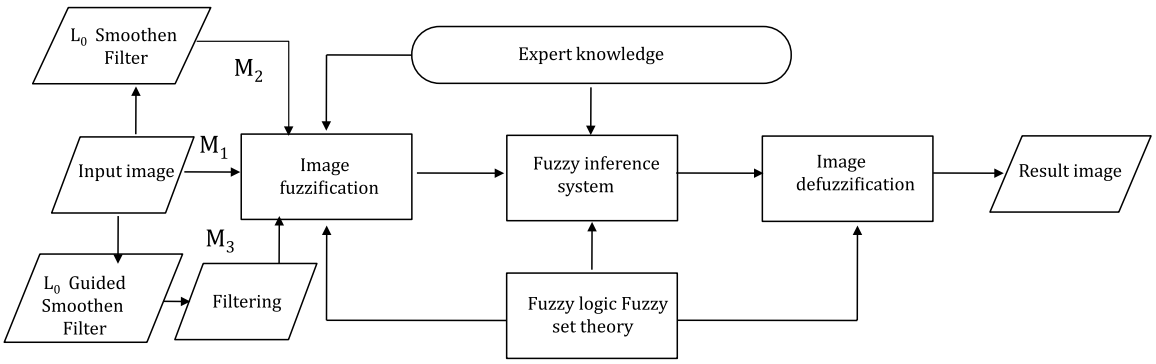


Fig. 3 Detailed architecture for edge detection mechanism

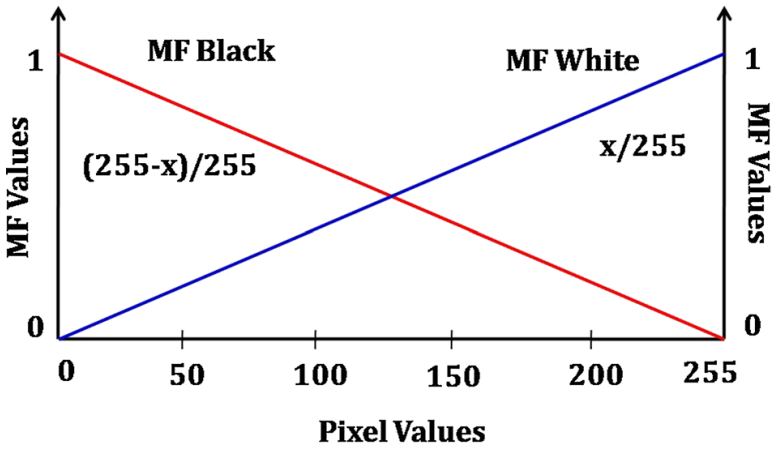


Fig. 4 Black and white pixel values membership functions

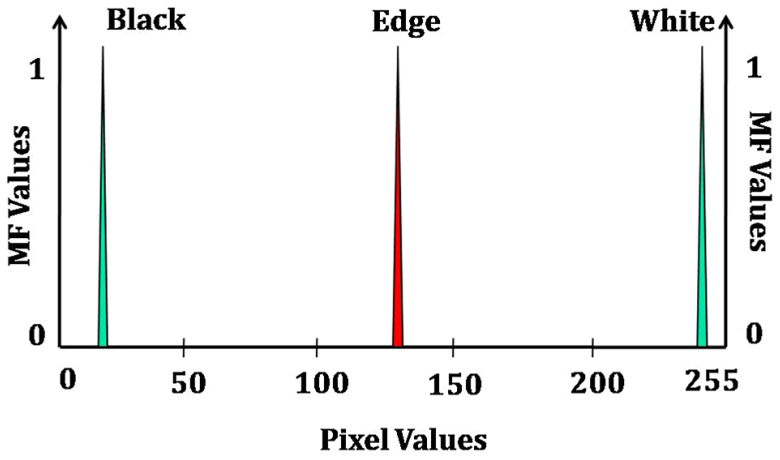


Fig. 5 Black, white, and edge outputs membership functions

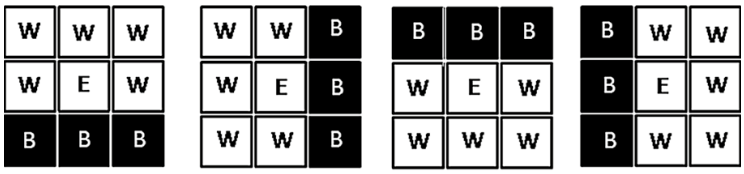
to note that for edge detection fuzzy logic is applied in all the methods. In the proposed method fuzzy logic is used for the edge detection. In fuzzy logic, first of all the input image is fuzzified by properly choosing input and output membership functions. Next fuzzy rules are derived, the membership function alonwith fuzzy set rules are fed to the Mamdani fuzzy inference engine, and the obtained output is de-fuzzified using the centroid method to obtain crisp values, which identifies a pixel as edge or non-edge pixel.

For both input and output, we employ a triangle membership function in this study. The input triangular function is defined as: $triangle(x;a,b,c) = \max\left(\min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right)$ as in Fig. 4. The output triangular membership functions for black, while and edge are very narrow as shown in Fig. 5.

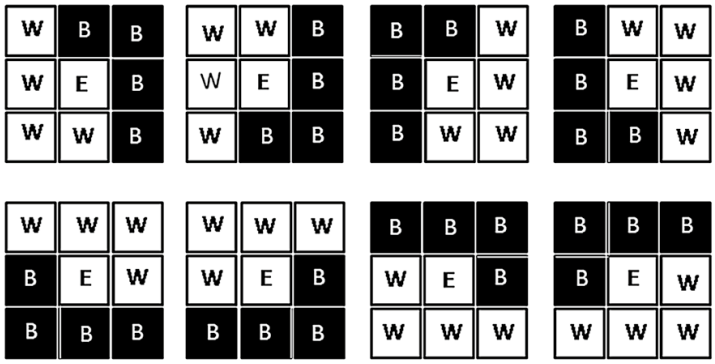
4.1 Fuzzy rules

A total of 30 rules are defined for edge classification, by considreing 3×3 windows as shown in Fig. 6. In the figure white pixel is deonted by '*W*' and black pixels by '*B*' and edge pixel by '*E*'. The rules are devised by conidering center pixel as edge pixel (*E*), and surrounding 8 pixels are varied.

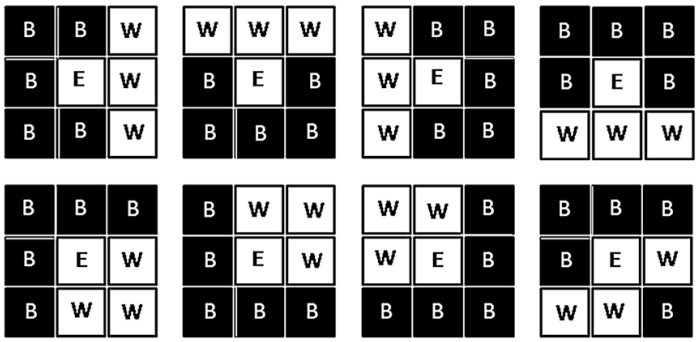
Rules with 3 B and 5 W pixels in neighbourhood of edge pixel (E)



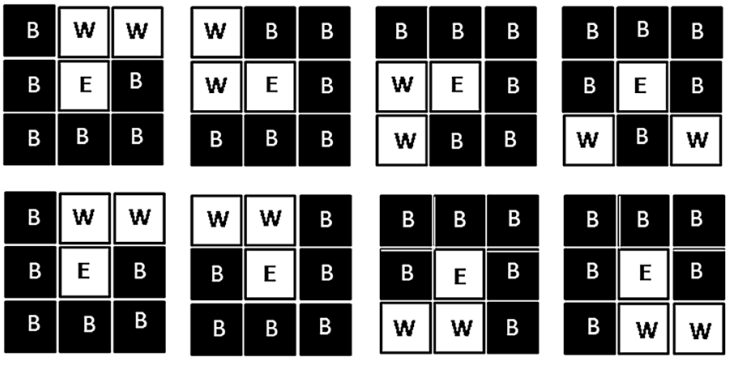
Rules with 4 B and 4 W pixels in neighbourhood of edge pixel (E)



Rules with 5 B and 3 W pixels in neighbourhood of edge pixel (E)



Rules with 6 B and 2 W pixels in neighbourhood of edge pixel (E)



Rules with 8 B or 8 W pixels in neighbourhood of non-edge pixel (NE)

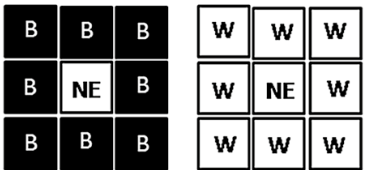


Fig. 6 Fuzzy rules for the classification of edge and non-edge pixels

5 Results

5.1 Performance Measures

It's difficult to design a generalized edge detection assessment since during edge detection mechanism some edges are falsely rejected, some are incorrectly accepted, and some change their positions. As a result, established metrics such as mean square error (MSE) and peak signal-to-noise ratio (PSNR) are ineffective. In this paper four measures Pratt's Figure of Merit [43], Structure Similarity Image Metrics [43], F-Score [43] and Hausdorff Distance [43] is considered.

5.1.1 Pratt's Figure of Merit (FoM)

By quantifying the displacement (d) of edge points identified from an ideal edge (I_A), the Pratt's Figure of Merit analyses edge location accuracy in edge detected images (I_B), in contrast to ground truth images. The Figure of Merit is characterized by

$$FoM = \frac{1}{\max(I_A, I_B)} \sum_{i=1}^{I_B} \frac{9}{\left(9 + d_i^2\right)} \tag{17}$$

The value of FoM lies between 0 and 1, when $d = 0$, FoM is 1, thus as FoM approaches 1, more accurate will be edge detection mechanism.

5.1.2 Structure Similarity Image Metrics (SSIM)

SSIM tries to estimate the changes in brightness, contrast, and structure in an image. The SSIM is given by

$$SSIM(A, B) = \frac{\left(2\mu_A\mu_B + k_1\right)(2\sigma_{AB} + k_2)}{\left(\mu_A^2 + \mu_B^2 + k_1\right)\left(\sigma_A^2 + \sigma_B^2 + k_2\right)} \tag{18}$$

In above equation, mean, variance and cross-correlation of the images A and B is represented by μ_A, μ_B,σ_A^2, σ_B^2 and σ_{AB} respectively, and rest terms are constants.

5.1.3 Hausdorff Distance (HoD)

The Hausdorff distance between two images $I_A = \{\alpha_1,\ldots,\alpha_n\}$ and $I_B = \{\beta_1,\ldots,\beta_n\}$, is given by

$$H(I_A, I_B) = \max(d(I_A, I_B), d(I_B, I_A))$$

where

$$d(I_A, I_B) = \max_{\alpha \in I_A} \min_{\alpha \in I_B} ||\alpha - \beta|| \tag{19}$$

The function $d(I_A, I_B)$ is the directed Hausdorff distance from I_A to I_B.

5.1.4 F-Score

In binary classification, the F-measure is a test of accuracy. The test score is based on both precession and recall. F has a maximum value of 1 and a minimum value of 0. It is the harmonic mean of precession and recall with equal weightage. Considering characteristic matrix (Fig. 7) the precession and recall are defined as

$$Precision\left(P_r\right) = \frac{True\ Positive}{(True\ Positive + False\ Positive)} \quad \text{and}$$

$$Recall\left(R_c\right) = \frac{True\ Positive}{(True\ Positive + False\ Negative)}$$

5.2 Comparative Results (Proposed Methods)

However, to prove usefulness of proposed method and to cover wide varieties of experiments results are presented on a single image considered from Berkley Segmentation Database (BSD) [46]. Finally, in the comparative analysis results are presented using both BSD and USC-SIPI Image Database is considered [47].

5.2.1 Results M_1 Method

In Fig. 8, results for edge detection are shown, using fuzzy logic only. It is clear from the figure that in edge detected images most of the edges are correctly detected with some more edges are falsely detected specially in Fig. 8c where at corners falsely detected edges can be easily seen.

5.2.2 Results M_2 Method

The falsely detected edges can be suppressed using L_0 smoothen filtering, but it should be kept in mind that more and more smoothness may leads to false rejection of edges. Therefore degree of smoothness plays an important role in detected edges. The chosen parameters for smoothening are $\beta_0 = 2\lambda$, $\beta_{\max} = 100{,}000$, $k = 2$ and λ varies (0.05–0.2).

Fig. 7 Characteristic matrix

	p	n
Y	True Positives	False Positives
N	True Negatives	False Negatives
Column Totals	P	N

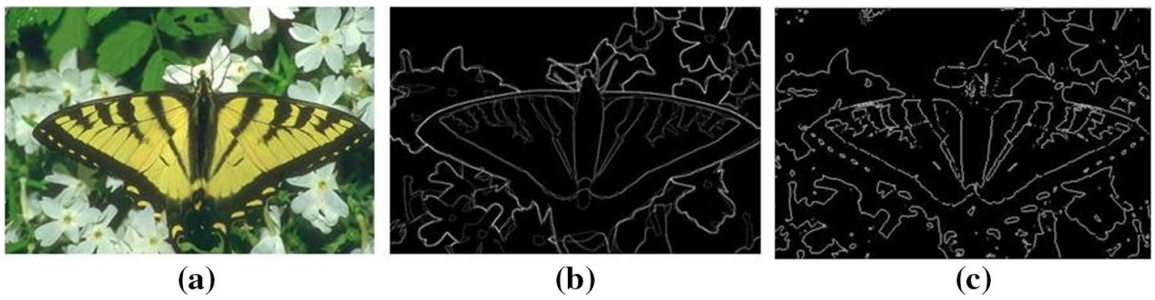


Fig. 8 **a** Original image, **b** ground truth, **c** detected edges using fuzzy logic

In Fig. 9, on the top row smoothen images are shown, using various level of degree of smoothness. As L_0 is edge preserving filter therefore even in smoothened prominent edges are preserved. In bottom row results for edge detection using fuzzy logic is shown on the smoothen images. It is clear from the figure that as smoothness increases, the falsely detected edges are reduced while prominent edges are still preserved.

5.2.3 Results M_3 Method

In Fig. 10, on the top row smoothen images are shown, using various level of degree of smoothness while considering guided filtering. As guided L_0 is edge preserving filter therefore even in smoothened prominent edges are preserved. On the bottom row edge detected images are shown. Therefore, by varying the degree of smoothness and applied methods (M_2 and M_3) the detected edges can be controlled. This is the main advantage of the proposed method where detected edges can be controlled.

In Fig. 11, comparison of various edge detection methods is shown, where in Fig. 11 ground truth image is shown with marked white areas where notable changes takes place in different methods. In Sobel method, edges in circular mark region are not properly detected. In case of Canny method a large number of falsely detected edges are found. In fuzzy method obtained results are well in agreement with ground truth image, but variation in rectangular mark region can be seen, along with some more falsely accepted

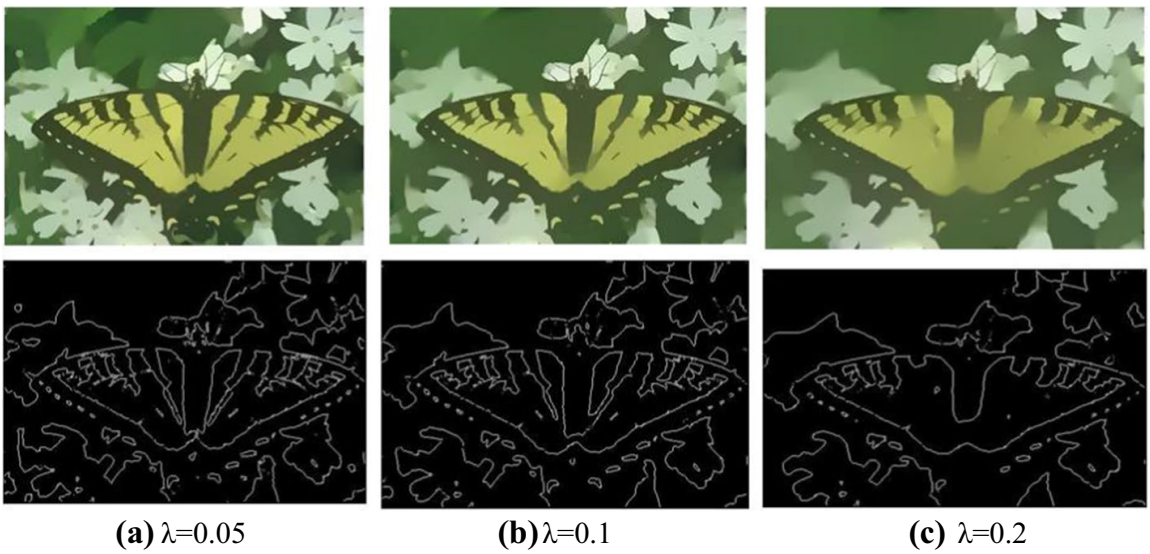


Fig. 9 Smoothen images using L_0 smoothing filter

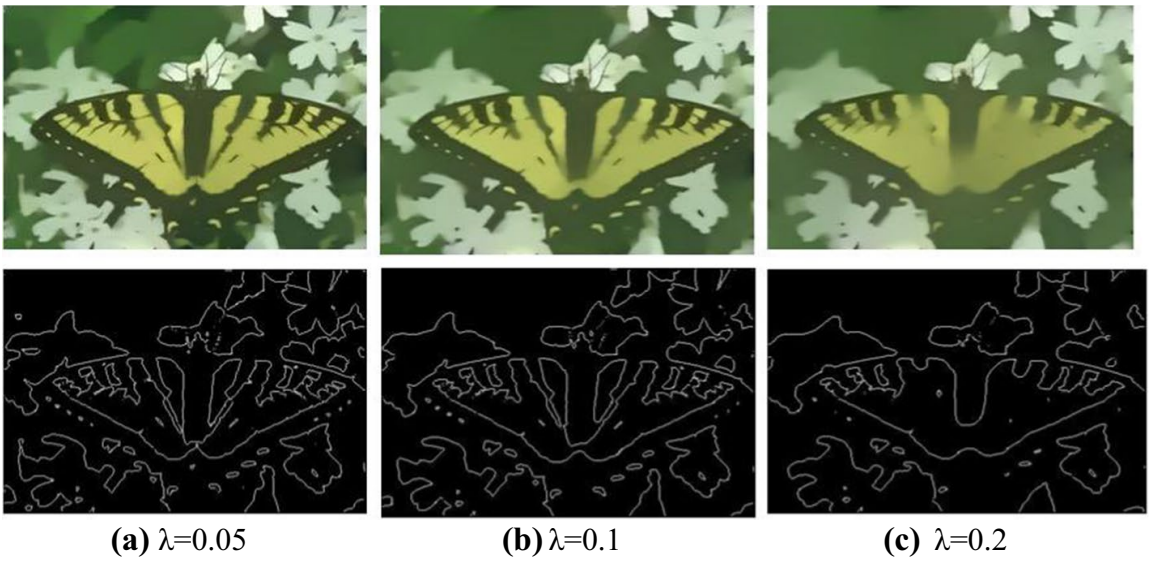


Fig. 10 Detected Edges in smoothen images using guided L_0 smoothing filter

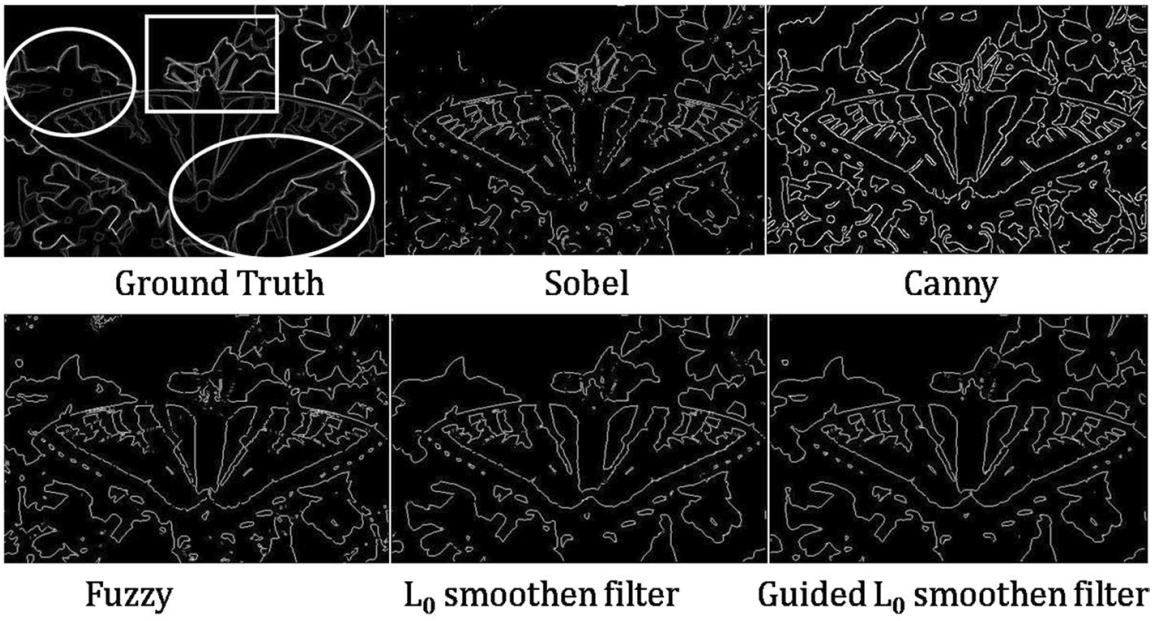


Fig. 11 Comparison of various edge detection methods

edges across the image. For L_0 smoothen filter and fuzzy method obtained results are well in agreement with ground truth image and variation is small, and finally for guided L_0 smoothen filter and fuzzy method, here again results are very much similar to L_0 smoothen filter and fuzzy method with minor inclusion of false edges in circular region, however oval mark section is best detected. It is clear from the figures that the variation is so small it is difficult to judge from naked eyes, therefore to judge the performance of the method three performance measure as discussed above are used and obtained results are shown in Figs. 12, 13 and 14.

In Figs. 12, 13 and 14, results for various edge detection techniques under three performance metrics are shown. Ideally FoM, SSIM should be one and HoD should be equal to zero. It is clear from the tables that FoM is lowest for canny edge detection, i.e., more pixels shift their position in comparison to other considered technique. In

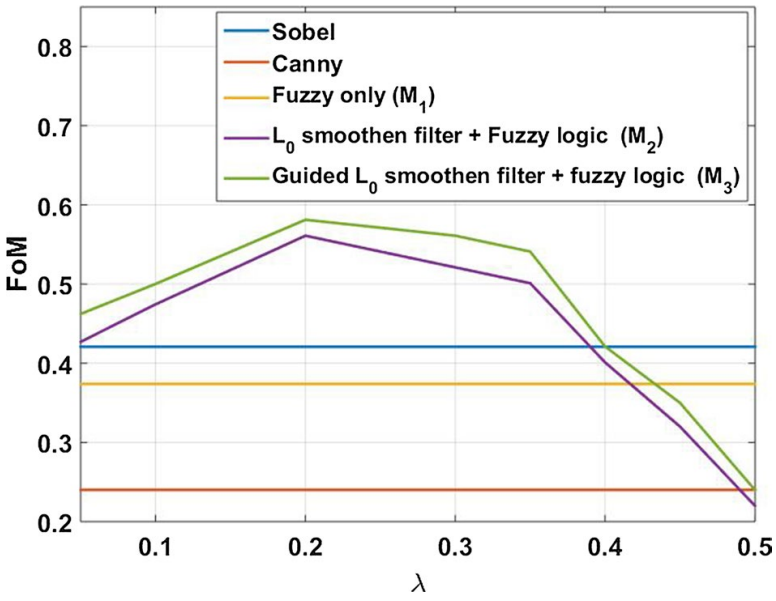


Fig. 12 Comparison of various edge detection methods (FoM vs. λ)

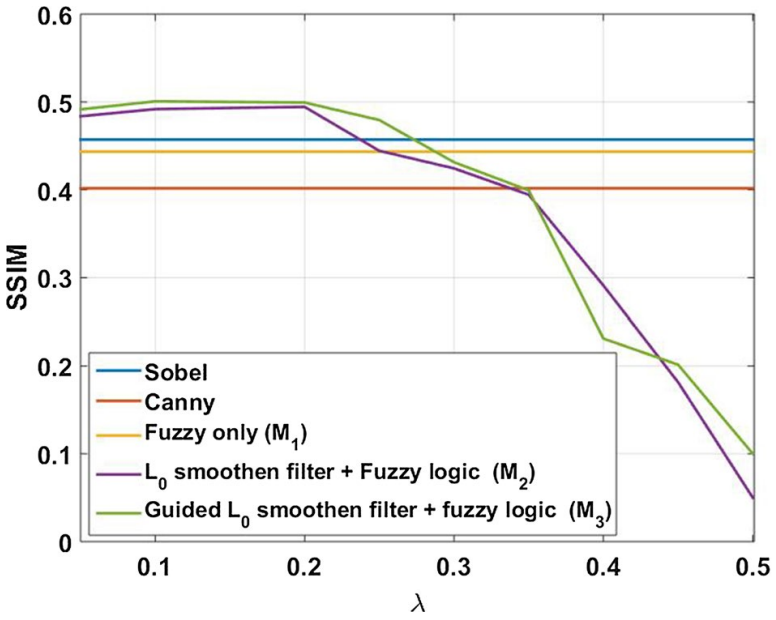


Fig. 13 Edge detection methods comparsions (SSIM vs. λ)

Fig. 12, FoM vs. λ is shown, It is evident that best FoM is obtained under Guided L_0 smoothen filter + fuzzy logic case. It is also observable that as degree of smoothness increases FoM increases up to a limit thereafter it starts to decreases. The best value of FoM is obtained for $\lambda = 0.2$. In Fig. 13, SSIM vs. λ is shown, again SSIM is comparatively better, for Guided L_0 smoothen filter + fuzzy logic case, but first it increases with degree of smoothness, thereafter SSIM decreases. This is obvious more smoothening will lead to structural modifications. In the considered cases, best SSIM is obtained for degree of smoothness of 0.1. In the Fig. 14, Hausdorff Distance *is* plotted, for best case average minimum distance is 3.30. Again here with increases in λ, HoD first decreases and then increases. These results clearly reveal that Guided L_0 smoothen filter + fuzzy logic provides better results for edge detection.

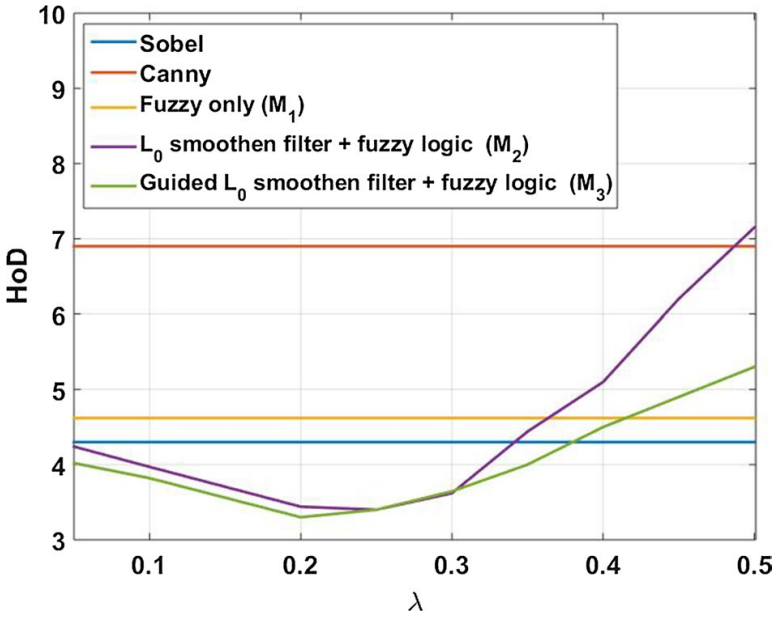


Fig. 14 Edge detection methods comparsions (HoD vs. λ)

5.3 Comparative with Recent Methods

In order to compare our method with more recent edge detection techniques, results are compared with Gonzalez et al., work where edges are detected using sobel and type-2 fuzzy logic method [48]. Results were tested on more than 100 images, and a few reproduced results are shown in Fig. 15.

In Fig. 15, four rows and four columns are shown, first column shows original image, and in second column results are shown for canny edge detection. In column 3 and 4 results are shown for Gonzalez et al. and proposed method respectively. In Canny method large numbers of false edges are accepted leads to the lowering of F-score. It is also observable that when intensity difference is less Canny method fail to detect edges (fourth row second column). Our method is comparable to Gonzalez et al., methods with less number of falsely accepted edges (second row and second, third columns).

In Table 1, notable and recently proposed methods are compared in terms of F-score. For classical methods canny and sobel F-measure is 0.49 and 0.40 respectively. The recently proposed learning based methods have F-score ranging from 0.63 to 0.78. A new kernel based method with singular value decomposition [42] has F-score as high as 0.83. Our proposed methods, L_0 smoothen filter + fuzzy logic (M_2) and Guided L_0 smoothen filter + fuzzy logic (M_3) attain F-score of 0.82 and 0.848 respectively.

The simulation is done MATLAB^R 2015a, on a syatem with quadcore processor, 4 GB RAM, the rum time for fuzzy logic method (M_1) is 3.67 s, while for L_0 smoothen filter + fuzzy logic (M_2) the runtime is 5.69 s. Finally, for Guided L_0 smoothen filter + fuzzy logic (M_3), the runtime is 7.43 s. Therefore, the proposed edge detection mechansim is fast enough that it can be applied to real time problems.

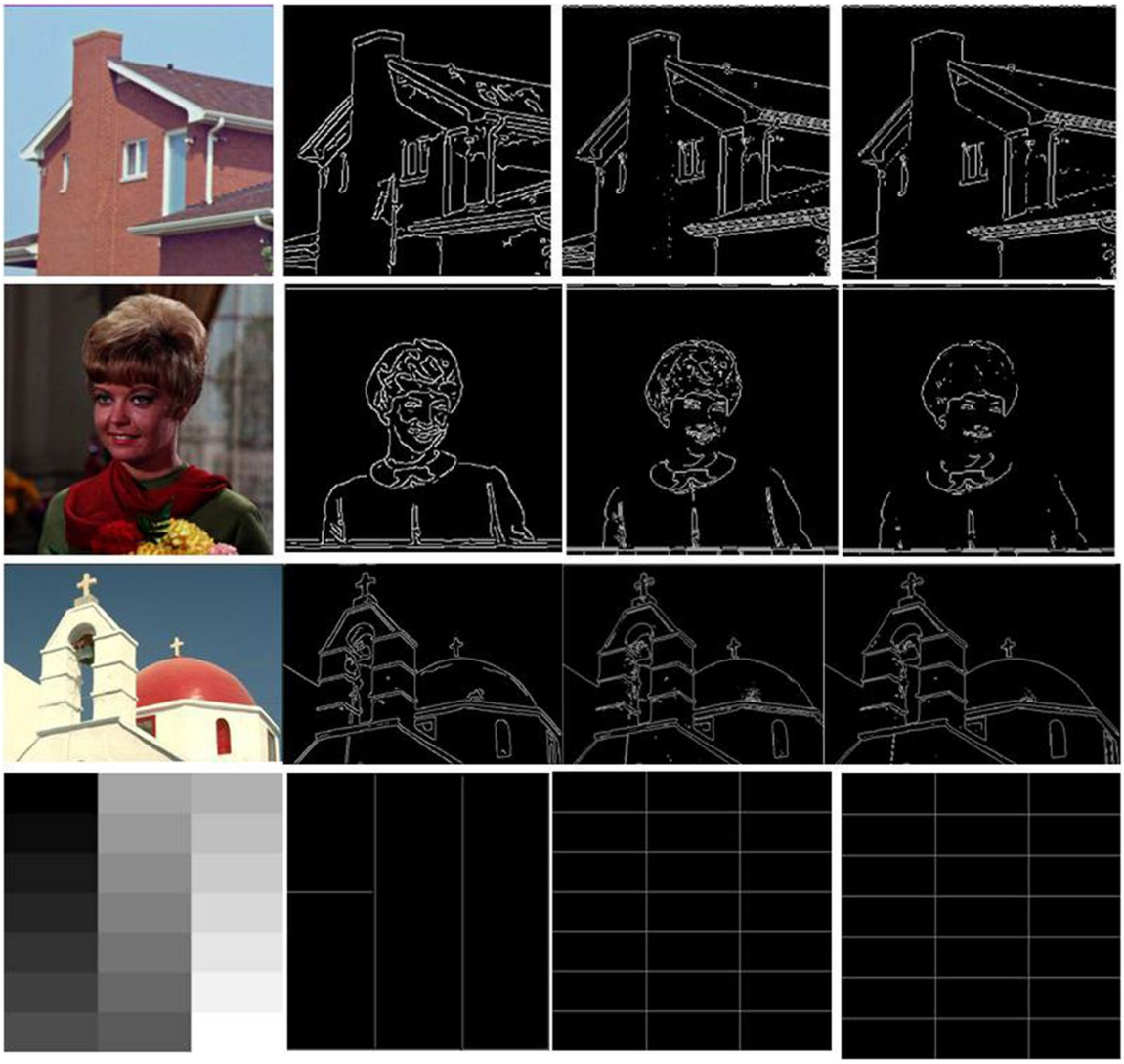


Fig. 15 Each row left to right: original, Canny, Gonzalez, et al., and proposed

Table 1 Comparison of notable edge detection methods

Methods	Year	F-measure
Canny [5]	(1996)	0.49
Sobel [49]	(2009)	0.40
BEL [50]	(2006)	0.63
gPb [12]	(2011)	0.71
Sketch Token [51]	(2013)	0.73
Structure Forest [52]	(2013)	0.71
Holistically-nested edge detection [53]	(2015)	0.78
Gonzalez et al. [48]	(2018)	0.83
Akshi et al. [19]	(2020)	0.79
Akshi et al. [20]	(2020)	0.81
Fuzzy only (M_1)	(2021)	0.77
L_0 smoothen filter + fuzzy logic (M_2)	(2021)	0.82
Guided L_0 smoothen filter + fuzzy logic (M_3)	(2021)	0.848

6 Conclusions

Edge detection has been an important area of research from past many years due to its utility in many fields like: image segmentation, medical, forensic and defense applications. In past years various methods based on kernels and soft computing based has been proposed, but they are all dependent on some kind of threshold mechanism and suffer from false acceptance/rejection. To deal such issues, in this paper a fuzzy based edge detection mechanism is proposed where edges are controlled using smoothen filters. In this paper two types of smoothen filters L_0 smoothen and guided L_0 smoothen filters are discussed, and using these filters prominent edges can be preserved and thus making edge detection more efficient. The performance of the edge detection methods are compared in terms of FoM, SSIM, F-measure and HoD, and it has been found that Guided L_0 smoothen filter + fuzzy logic produces better results. It is also found that smoothen should be done carefully and it should be within limit to obtain better results otherwise SSIM slips down and image quality goes down.

References

1. Petrou, M., & Petrou, C. (2010). *Image processing: The fundamentals*. Wiley.
2. Bovik, A. C. (2010). *Handbook of image and video processing*. Academic Press.
3. Poobathy, D., & Chezian, R. M. (2014). Edge detection operators: Peak signal to noise ratio based comparison. *International Journal of Image, Graphics and Signal Processing, 1*(10), 55–61.
4. Roberts, L. G. (1965). Machine perception of three-dimensional solids. In J. T. T ippett (Ed.) *Optical and electro-optical information processing* (Ch. 9, pp. 159–197). MIT Press.
5. Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 6*, 679–698.
6. Genming, C., & Baozong, Y. (1989). A new edge detector with thinning and noise resisting abilities. *Journal of Electronics, 6*(4), 314–319.
7. Matthews, J. (2002). *An introduction to edge detection: The sobel edge detector*. http://www.generation5.org/content/2002/im01.asp
8. Gonzalez, R. C., & Woods, R. E. (2002). *Digital image processing* (2nd ed.). Prentice Hall.
9. Grigorescu, C., Petkov, N., & Westenberg, M. A. (2003). Con-tour detection based on nonclassical receptive field inhibition. *IEEE Transactions on Image Processing, 12*(7), 729–739.
10. Ganin, Y., & Lempitsky, V. (2014). n4fields: Neural network nearest neighbor fields for image transforms. In *Asian conference on computer vision* (pp. 536–551). Springer.
11. Mairal, J., Leordeanu, M., Bach, F., Hebert, M., & Ponce, J. Discriminative sparse image models for class-specific edge detection and image interpretation. In *European conference on computer vision* (pp. 43–56). Springer.
12. Arbelaez, P., Maire, M., Fowlkes, C., & Malik, J. (2011). Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 33*(5), 898–916.
13. Xiaofeng, R., & Bo, L. (2012). Discriminatively trained sparse code gradients for contour detection. In *Advances in neural information processing systems* (pp. 584–592).
14. Dolĺar, P., & Zitnick, C. L. (2015). Fast edge detection using structured forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 37*(8), 1558–1570.
15. Verma, O. P., & Sharma, R. (2011). An optimal edge detection using universal law of gravity and ant colony algorithm. In *Proc. information and communication technologies (WICT)* (pp. 507–511).
16. Naumenko, A., Lukin, V., & Egiazarian, K. (2012). SAR-image edge detection using artificial neural network. In *Proc. international conference on mathematical methods in electromagnetic theory (MMET)* (pp. 508–512).
17. Setayesh, M., Zhang, M., & Johnston, M. (2012). Effects of static and dynamic topologies in particle swarm optimisation for edge detection in noisy images. In *Proc. evolutionary computation (CEC)* (pp. 1–8).

18. Gonzalez, C. I., Castro, J. R., Melin, P., & Castillo, O. (2015). Cuckoo search algorithm for the optimization of type-2 fuzzy image edge detection systems. In *2015 IEEE congress on evolutionary computation (CEC)* (pp. 449–455). IEEE.
19. Kumar, A., & Raheja, S. (2020). Edge detection using guided image filtering and ant colony optimization. In *The international conference on recent innovations in computing* (pp. 319–330). Springer.
20. Kumar, A., & Raheja, S. (2020). Edge detection using guided image filtering and enhanced ant colony optimization. *Procedia Computer Science, 173*, 8–17.
21. Bertasius, G., Shi, J., & Torresani, L. (2015). Deepedge: A multi-scale bifurcated deep network for top-down contour detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4380–4389).
22. Xie, S., & Tu, Z. (2017). Holistically-nested edge detection. *International Journal of Computer Vision, 125*(1–3), 3–18.
23. Wang, Y., Zhao, X., & Huang, K. (2017). Deep crisp boundaries. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3892–3900).
24. Liu, Y., Cheng, M.-M., Hu, X., Wang, K., & Bai, X. (2017). Richer convolutional features for edge detection. In *Computer vision and pattern recognition (CVPR), 2017 IEEE Conference on* (pp. 5872–5881). IEEE.
25. Marmanis, D., Schindler, K., Wegner, J. D., Galliani, S., Datcu, M., & Stilla, U. (2018). Classification with an edge: Improving semantic image segmentation with boundary detection. *ISPRS Journal of Photogrammetry and Remote Sensing, 135*, 158–172.
26. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
27. Shen, W., Wang, X., Wang, Y., Bai, X., & Zhang, Z. (2015). Deepcontour: A deep convolutional feature learned by positive-sharing loss for contour detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3982–3991).
28. Bertasius, G., Shi, J., & Torresani, L. (2015). Deepedge: A multiscale bifurcated deep network for top-down contour detection. In *CVPR*.
29. Hwang, J.-J., & Liu, T.-L. (2015). Pixel-wise deep learning for contour detection. In *ICLR*.
30. Zhang, W., Zhao, Y., Breckon, T. P., & Chen, L. (2017). Noise robust image edge detection based upon the automatic anisotropic Gaussian kernels. *Pattern Recognition, 63*, 193–205.
31. Farbod, M., Akbarizadeh, G., Kosarian, A., & Rangzan, K. (2018). Optimized fuzzy cellular automata for synthetic aperture radar image edge detection. *Journal of Electronic Imaging, 27*(1), 013030.
32. Ren, H., Zhao, S., & Gruska, J. (2018). Edge detection based on single-pixel imaging. *Optics express, 26*(5), 5501–5511.
33. Cosgrove, C., & Yuille, A. L. (2019). Adversarial examples for edge detection: They exist, and they transfer. arXiv preprint arXiv:1906.0033.
34. Kuo Y. H., Lee C. S., & Liu, C. C. (1997). A new fuzzy edge detection method for image enhancement. In *Proc. 6th IEEE international conference* (Vol. 2, pp. 1069–1074).
35. Russo, F. (1998). Edge detection in noisy images using fuzzy reasoning. In *Proc. instrumentation and measurement technology conference* (pp. 369–372). IMTC/98.1.
36. El-Khamy S. E., Lotfy M., & El-Yamany N. (2000). A modified fuzzy Sobel edge detector. In *Proc. radio science conference* (pp. C32-1).
37. Kim, D. S., Lee, W. H., & Kweon, I. S. (2004). Automatic edge detection using 3× 3 ideal binary pixel patterns and fuzzy-based edge thresholding. *Pattern Recognition Letters, 25*(1), 101–106.
38. Alshennawy, A. A., & Aly, A. A. (2009). Edge detection in digital images using fuzzy logic technique. *World Academy of Science, Engineering and Technology, 24*(51), 1781–1786.
39. Tab, F. A., Shahryari, O. K. (2009). Fuzzy edge detection based on pixel's gradient and standard deviation values. In *Proc. computer science and information technology, IMCSIT'09*, (pp. 7–10).
40. Kaur, E. K., Mutenja, V., & Gill, E. I. (2010). Fuzzy logic based image edge detection algorithm in MATLAB. *International Journal of Computer Applications, 1*(22), 55–58.
41. Hsu, C. H., & Juang, C. F. (2011). Evolutionary robot wall-following control using type-2 fuzzy controller with species-DE-activated continuous ACO. *IEEE Transactions on Fuzzy Systems, 21*(1), 100–112.
42. Chen, S. M., Chang, Y. C., & Pan, J. S. (2013). Fuzzy rules interpolation for sparse fuzzy rule-based systems based on interval type-2 Gaussian fuzzy sets and genetic algorithms. *IEEE Transactions on Fuzzy Systems, 21*(3), 412–425.
43. Raheja, S., & Kumar, A. (2021). Edge detection based on type-1 fuzzy logic and guided smoothening. *Evolving Systems, 12*, 447–462.
44. Xu, L., Lu, C., Xu, Y., & Jia, J. (2011). Image smoothing via L_0 gradient minimization. *ACM Transactions on Graphics, 30*(6), 174.

45. Ding, X., Chen, L., Zheng, X., Huang, Y., & Zeng, D. (2016). Single image rain and snow removal via guided L_0 smoothing filter. *Multimedia Tools and Applications, 75*(5), 2697–2771.
46. https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/BSDS300/html/dataset/images.html
47. http://sipi.usc.edu/database/database.php
48. Gonzalez, C. I., Melin, P., Castro, J. R., Mendoza, O., & Castillo, O. (2016). An improved sobel edge detection method based on generalized type-2 fuzzy logic. *Soft Computing, 20*(2), 773–784.
49. Vincent, O. R., & Folorunso, O. (2009). A descriptive algorithm for sobel image edge detection. *Proceedings of Informing Science & IT Education Conference, 40*, 97–107.
50. Dollar, P., Tu, Z., & Belongie, S. (2006). Supervised learning of edges and object boundaries. In *Computer vision and pattern recognition, 2006 IEEE computer society conference on* (Vol. 2, pp 1964–1971).
51. Lim, J. J., Zitnick, C. L., & Dollár, P. (2013). Sketch tokens: A learned mid-level representation for contour and object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3158–3165).
52. Dollár, P., & Zitnick, C. L. (2013). Structured forests for fast edge detection. In *Proceedings of the IEEE international conference on computer vision* (pp. 1841–1848).
53. Xie, S., & Tu, Z. (2015). Holistically-nested edge detection. In *Proceedings of the IEEE international conference on computer vision* (pp. 1395–1403).

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



Dr. Akshi Kumar is an Associate Professor at the Department of Information Technology, Netaji Subhas University of Technology, Delhi, India. She is a Post-doc from Federal Institute of Education, Science and Technology of Ceará, Fortaleza, Brazil and PhD from Faculty of Technology, University of Delhi, India. She was an Assistant Professor at the Delhi Technological University (formerly, Delhi College of Engineering), Delhi, India and has a total teaching experience of 16 years.My core teaching and research interests are in the domains of Web Technologies and Intelligent Systems. I conduct research in the design, analysis, and applications of Web. I also use techniques from Machine Learning, Soft Computing, Natural Language Processing, Probability and Information Theory in my research. I am ardently involved in projects on Sentiment Analysis in Web 2.0 Technologies. My focus lies in the area of investigating paradigms, devising techniques and building tools that perform sentiment analysis on social media (Twitter, Facebook, Blog, Quora) with suitable practical applications for business and government intelligence.



Mr. Sahil Raheja is a research scholar in the department of Information and Technology, Delhi Technological University. He has completed his M.Tech in Information systems from Delhi Technological University in 2012. He also worked as an assistant professor in School of computing science and engineering at Galgotias University, Greater Noida, India and has a total teaching experience of 10 years. His area of interests are Image and Video Processing, Machine Vision, Soft computing.

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

Efficient Methods to Improve the Performance of Supervised Learning Models

Mohit Kumar
Department of Electrical Engineering
Delhi Technological University
Delhi, India
mohitkumar451998@gmail.com

Pragya Yadav
Department of Electrical Engineering
Delhi Technological University
Delhi, India
pragyayadav011@gmail.com

Harman Singh
Department of Electrical Engineering
Delhi Technological University
Delhi, India
harman8911@gmail.com

Ankita Arora
Department of Electrical Engineering
Delhi Technological University
Delhi, India
arora.ankita09@gmail.com

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498387

Abstract: **Supervised Learning is defined as training a model with input data that includes the result itself. There are large number of supervised learning algorithms and great number of models. Each model has its own merits and demerits and performs differently. There are many data preprocessing techniques and hence the combination of several data preprocessing techniques can increase the performance of the present supervised learning models. Primary data can not be fed directly to the learning model because it can hold a lot of noise. It needs to be preprocessed using various data preprocessing techniques. We have analysed and compared different data preprocessing techniques and their combinations. Comparison is done using various performance metrics and the combination of different data preprocessing is applied to different model. We have done categorical data handling, missing value treatment, feature scaling and feature extraction as the data preprocessing steps. Through the comparison, we found which technique is better for which type of models. We used the California census data for our study.**

Keywords: Supervised Learning, Feature Scaling, Data Preprocessing, Feature Extraction.

I. INTRODUCTION

The performance of a supervised learning model indicates how beneficial predictions are made by the trained model. Supervised Machine Learning first takes the instances along with the labels from the outer world to produce some hypothesis and then predict the labels for future instances. One cannot directly use raw data to train a model as it contains noise and irrelevant information. The quality of data determines the success of a machine learning model and is of utmost importance. Data preprocessing is required to process the raw data and extract relevant information from that which can be fed to the machine learning model. Various supervised machine learning models have been introduced and compared [1]. It is required to make advances in the data preprocessing techniques to improve the performance of existing machine learning models. Data preprocessing includes missing value treatment, scaling of attributes, handling categorical attributes, extraction of relevant features, feature engineering, data compression, combination of attributes, etc [2].

One of the solutions for treating missing values is to either discard that attribute which has missing values or that instance which contains missing value. In the real world, there are so many missing values and discarding these values can result in a high loss of information. If we discard the tuples or the attributes which have missing values then we can lose significant amount of information. In the real world, the data which is collected contains a lot of missing values and almost each and every tuple as well as attribute contains missing values. So instead of discarding the tuples or columns we can treat the missing value. Batista Gustavo et al. [3] treated the missing values using an imputation method based on k-nearest neighbour algorithm and showed that it outperforms the internal methods used by C4.5 and CN2 for treating missing values. C4.5 uses probabilistic approach to compute the missing values. k-nearest neighbour algorithm performs better than this probabilistic approach. Stekhoven Daniel et al. [4] proposed an iterative imputation method MissForest based on random trees. It works better than the conventional methods for nonlinear data. Median or mean imputation cannot be used for nonlinear data. Different out of the box methods need to be used for non linear data imputation. Karabulut Esra Mahsereci et al. [5] showed that the performance of various models improved after the feature selection is applied to the dataset. Feature selection technique takes those features into account which are important for making predictions. Significant improvement in the accuracy of Naive Bayes, Artificial Neural Network and J48 is observed after applying feature selection methods. Feature selection reduces the noisy attributes and hence improve the performance of black box machine learning algorithm such as neural network. Jayaram M.A. et al. [6] has selected the feature subset using the wrapper approach. He used Genetic algorithm as a random search technique with various induction algorithms and the results are validated using classifiers. He concludes that there is not a specific wrapper method which is better than the other but feature selection improves the classification accuracy. Wrapper methods create a feature subset and select the subset from the search space having best performance on the dataset. Chandrashekar Girish et al. [7] uses feature selection techniques and improved the performance for fault prediction analysis for Fault Mode Data. More data is always not good for machine learning models. The curse of dimensionality overfits the model on the present dataset and hence deteriorate the performance of machine learning models. Hence the more

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:53:08 UTC from IEEE Xplore. Restrictions apply.

irrelevant data is not useful for machine learning algorithms. Only relevant data is required so that the algorithm can outperform. Lazar Cosmin et al. [8] surveyed filter feature selection methods for valuable feature discovery in differentially expressed genes. Filter methods are independent of any machine learning techniques. Using filter methods, features are selected on the basis of various statistical tests performed on the data. Jerez Jose M. et al. [9] applied six imputation methods based on statistical methods such as mean imputation, hot-deck, machine learning techniques etc. and compared the results obtained from listwise deletion. It is concluded that the machine learning imputation techniques performed better. Heaton Jeff [10] has analysed feature engineering methods producing ten different features and showed that random forest and gradient boosting machines perform well for the same type of engineered features. It also showed that ensemble methods perform better than the individual methods.

Our approach uses some data preprocessing techniques to improve the performance of supervised learning models. We use two categorical data handling methods, two missing values treatment methods, two feature scaling methods and two feature extraction methods. In total we use 16 different combinations of data preprocessing techniques and then compare the performance of different models using these techniques. We use four different regression models to compare the results as Linear Regressor, Decision Tree Regressor, Random Forest Regressor, Support Vector Regressor and Gradient Boosting Regressor. We perform a comparative study on which techniques are effective for the performance of supervised learning models.

We calculate various performance metrics for each regressor and then based on that compared each data preprocessing techniques. Each model has a different capability. So each should perform differently than the other model for each data preprocessing technique. Each regressor model has their own advantages and disadvantages so they will perform differently on the dataset, more specifically on the data which lie on the boundaries. We have used linear models, tree models as well as ensemble models.

Hence our study provides the following advantages over the other works performed earlier:

1. Comparative study of different combinations of various data preprocessing techniques.

2. Combination of data preprocessing techniques which are better for which type of models are shown based on the values of different performance metrics.

The rest of the paper shows our approach followed by comparative results. The last section contains the conclusion of the paper.

II. Modelling of Supervised Learning Methods and Discussion Results

Our study aims to compare various data preprocessing approaches which can increase the performance of supervised learning models. We have implied the following data preprocessing techniques: categorical data handling, missing value treatment, data scaling and feature extraction. Each data preprocessing technique uses two types of approach. We get 16 different approaches in total for data preprocessing. Using these techniques data is preprocessed and then the results of various supervised learning models are compared using various performance metrics.

A. Categorical Data Handling

The data preprocessing starts with treating categorical data. Most of the machine learning libraries work on numeric data and cannot handle the categorical data. The categorical data is represented in some numeric form to make it usable for feeding the machine learning models. We propose two methods for treating categorical data: conversion into a number and dummy coding.

1) Conversion into a number converts the different levels: The categorical data of the specific attribute to unique numbers. We take the range of the numbers starting from including zero to the levels of that categorical attribute.

2) Dummy coding: It is commonly used to convert the categorical data into numeric data. Dummy variables are created for each level present in the categorical attribute. The dummy variable is marked as one if that level corresponds to the given instance else that dummy variable is marked as zero. Each of these methods converts the categorical data to numeric data. Using the first method the dimension remains the same as the old categorical attribute is replaced by the new numeric attribute but using the second method increases the dimension of the dataset as the old categorical attribute is replaced by new dummy variables.

B. Missing Value Treatment

Missing data can deteriorate the performance of a machine learning model. Missing data can be treated by deleting the attribute or the instances having missing value. Missing value can also be treated by filling the missing value by some method. In this approach we treat the missing values by two ways: Mean/Mode/Median imputation and KNN Imputation.

1) Mean Imputation: It replaces the missing values by calculating the mean or median or mode of all the non-missing values of the feature.

2). KNN Imputation: It imputes the missing value by finding the specified nearest instances having nonmissing values. The nearest instances are specified by the distance function. It can impute both qualitative and quantitative attributes. Attributes with multiple missing values can easily be treated. This algorithm is very time consuming.

C. Feature Scaling

Machine learning algorithms such as linear regression, etc. which uses gradient descent as an optimization approach. Features having the same range help the gradient descent to converge quickly at the minima. Distance based algorithms such as SVM, etc. also require scaling so that all features contribute equally. We use two approaches to scale the data: standardisation and normalisation.

1) Standardisation: It is a scaling technique in which the data is centred around the mean with a unit standard deviation. We obtain the mean μ and the standard deviation

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:53:08 UTC from IEEE Xplore. Restrictions apply.

σ of the feature values, the standardise the data X using the formula

$$X' = \frac{X - \mu}{\sigma}$$

2) Normalisation: It is a scaling technique in which the data is shifted and scaled to a range of [0,1]. We represent the minimum value as $xmin$ and the maximum value as $xmax$, then we normalise the data X using the formula

$$X' = \frac{X - x_{min}}{x_{max} - x_{min}}$$

Fit the scaler on the training data and not on the complete data or test data.

D. Feature Selection

The data consists of the both relevant and irrelevant information. Feature selection techniques select the features containing the most relevant information and discards those features which behave like noise. Here, we have used two feature selection methods: correlation based feature selection and recursive feature elimination.

1) Correlation: It defines the linear relationship between two variables. The correlation coefficient is calculated between the dependent and all other independent variables and then the specified number of features having highest correlation with the target feature are selected.

2) Recursive Feature Elimination: It removes the weakest feature until the specified number of features are reached. It tries to remove collinearity and dependencies which can exist in the model. Variable importance is determined by feature importance or correlation.

E. Prediction

After performing the proposed data preprocessing steps, to compare the results of our different data preprocessing approaches, we have selected five regressors namely Linear Regression, Decision Tree Regressor, Random Forest Regressor, Support Vector Regressor as well as ensemble models such as Gradient Boosting Regressor. All regressors predict the continuous value of the target variable.

1) Linear Regression: It fits a line y = mx + c to the given training data and predicts the test variable using this line. In case of multiple variables, a multidimensional plane fits to the training data and it is used for further prediction.

2) Decision Tree Regressor: It creates a tree where each node determines which type of instances fit that path. Each node uses an attribute to break it into further nodes. It predicts the value as the value corresponding to the leaf node where the instance reaches.

3) Random Forest Regressor: It is a collection of many decision tree regressors. The target value is predicted using all the decision trees and then averaged to get the final answer. It is also a kind of ensemble model.

4) Support Vector Regressor: It works the same as a support vector machine. It creates a hyperplane which maximises the margin and predicts the value for the target variable.

5) *Gradient Boosting Regressor*: It is an ensemble learning model. It adds the predictors sequentially by correcting its predecessor. When decision trees are used as predictors then it is called Gradient Boosting Regressor Trees.

All the regression models predict a continuous numeric value of the target variable using the value of independent variables.

III. Results And Table Discussion

All the models for this study have been done on a system with Ubuntu 18.04.3 LTS having 2.40 GHz Intel Core i5 processor and Python 3. To compare the results of various models using different combinations of data preprocessing methods, we have chosen the California Housing Prices dataset from the StatLib repository [11]. This dataset is based on data from the 1990 California census. The complete dataset has been splitted into ten parts to perform the hold one out validation and obtain more reliable results. For comparing the results of various models using different data preprocessing pipelines, we adopted five performance metrics namely mean squared error, mean absolute error, root mean squared error, r-square error and adjusted r square-error.

A. Comparison of data preprocessing with benchmark regressors

In this study, we have compared our data preprocessing methods using various supervised learning models such as Linear Regression, Decision Tree Regressor, Random Forest Regressor, Support Vector Regressor as well as ensemble models such as Gradient Boosting Regressor.

Table I, II, III, IV and V shows the results of various performance metrics applied to different algorithm using various combination of data preprocessing techniques. It can be seen that ensemble method as Random Forest Regressor and Gradient Boosting Regressor perform better that the linear regressor as well as decision tree regressor. Support Vector Regressor does not perform well over this dataset. But Gradient Boosting Regressor performs better than the Random Forest Regressor for every technique.

In case of Linear Regressor, combination of dummy coding, knn imputation, standardisation and correlation attribute selection works the best. In case of Decision Tree Regressor and Random Forest Regressor, combination of dummy coding, knn imputation, normalisation and correlation attribute selection works the best. In case of Support Vector Regressor, combination of dummy coding, mean imputation, normalisation and recursive feature elimination works the best. In case of Linear Regressor, combination of label coding, knn imputation, standardisation and recursive feature elimination works the best.

So comparing the various combination over different models we get a fair idea for which data preprocessing technique works for which type of algorithm.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:53:08 UTC from IEEE Xplore. Restrictions apply.

TABLE I: PERFORMANCE OF LINEAR REGRESSION

Technique	Mean Squared Error	Mean Absolute Error	Root mean squared error	R square error	Adjusted r square error
label_mean_std_corr	6374796738	59544.1084	78617.013	0.42276386	0.422596
label_mean_std_rfe	6394720219	59038.1103	79364.294	0.48672037	0.406739
label_mean_norm_corr	12735580803	87997.2882	111259.321	0.17090386	0.17892146
label_mean_norm_rfe	12956390284	80378.0392	118566.205	0.18574937	0.17595639
label_knn_std_corr	6374796738	59544.1084	78617.013	-0.422763860	-0.42317759
label_knn_std_rfe	6395762020	50283.2937	70365.365	-0.486292729	-0.40275089
label_knn_norm_corr	12735580803	87997.2882	111259.321	0.17090386	0.17066276
label_knn_norm_rfe	12947492028	80387.0376	115486.567	0.18464829	0.14049264
dummy_mean_std_corr	5863782984	49544.1084	68617.013	0.52276386	0.522596
dummy_mean_std_rfe	5803836120	40175.2930	60264.784	0.50275403	0.502796
dummy_mean_norm_corr	11739462820	77997.2882	101259.321	0.27090386	0.27892146
dummy_mean_norm_rfe	11947493740	725488.364	103547.494	0.20562938	0.27946259
dummy_knn_std_corr	5863782984	49544.1084	68617.013	-0.322763860	-0.32317759
dummy_knn_std_rfe	5803864828	40174.2784	69364.204	-0.302750572	-0.30576282
dummy_knn_norm_corr	11739462820	77997.2882	101259.321	0.27090386	0.27066276
dummy_knn_norm_rfe	11048482920	72964.8204	104769.706	0.20574374	0.27027484

TABLE II: PERFORMANCE OF DECISION TREE REGRESSOR

Technique	Mean Squared Error	Mean Absolute Error	Root mean squared error	R square error	Adjusted r square error
label_mean_std_corr	12339201277	77013.49	107476.443	-0.14253155	-0.07899738
label_mean_std_rfe	12750374048	77364.1944	104759.485	-0.17483947	-0.07027948
label_mean_norm_corr	11148537202	75451.0315	104415.264	0.032423615	0.036725373
label_mean_norm_rfe	11948363304	79365.3759	103894.485	0.030575927	0.030588595
label_knn_std_corr	11428723781	74597.2272	104807.292	0.085368362	0.085102391
label_knn_std_rfe	11963739203	79254.4759	107304.374	0.083694739	0.083936202
label_knn_norm_corr	11000346901	74999.5095	103695.604	0.040634007	0.040355026
label_knn_norm_rfe	11947495057	76374.2649	105303.374	0.049270282	0.049274947
dummy_mean_std_corr	11339205289	67013.49	97476.443	-0.04253155	-0.05899738
dummy_mean_std_rfe	11748463894	68364.2650	96204.384	-0.04959275	-0.05927492
dummy_mean_norm_corr	10148534209	65451.0315	94415.264	0.132423615	0.136725373
dummy_mean_norm_rfe	10746382937	69365.3850	96207.287	0.130583739	0.130559227
dummy_knn_std_corr	10428724135	64597.2272	94807.292	0.185368362	0.185102391
dummy_knn_std_rfe	10468339003	68254.5720	90364.384	0.189575938	0.180374937
dummy_knn_norm_corr	10000348430	64999.5095	93695.604	0.140634007	0.140355026
dummy_knn_norm_rfe	10369573648	67483.3884	92564.408	0.140686739	0.140472929

TABLE III: PERFORMANCE OF RANDOM FOREST REGRESSOR

Technique	Mean Squared Error	Mean Absolute Error	Root mean squared error	R square error	Adjusted r square error
label_mean_std_corr	7217492052	61608.234	82785.6461	0.327075547	0.32571246
label_mean_std_rfe	7203730227	60274.3943	86394.978	0.303840802	0.39750394
label_mean_norm_corr	6140788957	57345.6677	77035.005	-0.44506612	-0.44244165
label_mean_norm_rfe	6103783728	59387.2086	70265.985	-0.40284028	-0.48374739
label_knn_std_corr	7212618375	61634.3930	82509.522	-0.32512854	-0.32551388
label_knn_std_rfe	7293739479	60265.9254	84485.095	-0.39474284	-0.30484944
label_knn_norm_corr	6147323521	57445.4265	77081.524	-0.445107885	-0.44493084
label_knn_norm_rfe	6108023837	50275.9375	74385.474	-0.408840382	-0.44827494
dummy_mean_std_corr	6722992052	51608.234	72785.6461	0.427075547	0.42571246
dummy_mean_std_rfe	6793730273	56578.8562	72648.394	0.402840482	0.48364930
dummy_mean_norm_corr	5184088957	47345.6677	67035.005	-0.34506612	-0.34244165
dummy_mean_norm_rfe	5193830282	40987.3638	68376.305	-0.38475048	-0.39790303
dummy_knn_std_corr	6194018375	51634.3930	72509.522	-0.22512854	-0.22551388
dummy_knn_std_rfe	6193720284	59275.0275	76204.495	-0.20494928	-0.29474072
dummy_knn_norm_corr	5015823521	47445.4265	67081.524	-0.345107885	-0.34493084
dummy_knn_norm_rfe	5093749393	49867.3096	69376.384	-0.304830847	-0.37494839

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:53:08 UTC from IEEE Xplore. Restrictions apply.

TABLE IV: PERFORMANCE OF SUPPORT VECTOR REGRESSOR

Technique	Mean Squared Error	Mean Absolute Error	Root mean squared error	R square error	Adjusted r square error
label_mean_std_corr	14398637636	89667.7278	117299.214	-0.26571505	-0.26608311
label_mean_std_rfe	14769265025	85309.6397	118639.493	-0.26937103	-0.26917010
label_mean_norm_corr	14499708154	90041.4420	117733.478	0.27511489	0.274904101
label_mean_norm_rfe	14968539503	90365.7782	110375.798	0.27028457	0.270183820
label_knn_std_corr	14398637636	89667.7278	117299.214	0.265715046	0.265501518
label_knn_std_rfe	14970255036	80174.9479	110476.027	0.269207451	0.260193754
label_knn_norm_corr	14499708154	90042.4420	117733.478	0.275114895	0.2749041019
label_knn_norm_rfe	14067144067	97254.9408	110688.375	0.270294739	0.2701974529
dummy_mean_std_corr	13650157893	79667.7278	107299.214	-0.26571505	-0.26608311
dummy_mean_std_rfe	13960285304	70275.0578	100364.204	-0.26620674	-0.26927210
dummy_mean_norm_corr	13864190453	80041.4420	107733.478	0.27511489	0.274904101
dummy_mean_norm_rfe	13077249567	80285.9300	106394.047	0.27936505	0.270340370
dummy_knn_std_corr	13760165892	79667.7278	107299.214	0.265715046	0.265501518
dummy_knn_std_rfe	13086420867	70285.0275	108354.028	0.269364927	0.269103852
dummy_knn_norm_corr	13770158706	80042.4420	107733.478	0.275114895	0.2749041019
dummy_knn_norm_rfe	13029768257	80295.3839	109367.208	0.279372927	0.2708352910

TABLE V: PERFORMANCE OF GRADIENT BOOSTING REGRESSOR

Technique	Mean Squared Error	Mean Absolute Error	Root mean squared error	R square error	Adjusted r square error
label_mean_std_corr	5491005948	59667.7278	72754.428	0.503925384	0.503781128
label_mean_std_rfe	5380174837	58556.6167	71643.428	0.514936495	0.514936495
label_mean_norm_corr	6101700817	58117.1746	76286.799	-0.46293856	-0.46336398
label_mean_norm_rfe	6000600706	57006.0635	75175.799	-0.45184745	-0.45184745
label_knn_std_corr	5488026893	54186.7312	72738.095	-0.50412679	-0.504564188
label_knn_std_rfe	5377015782	53075.6201	71627.095	-0.50301568	-0.50301568
label_knn_norm_corr	6101696465	58117.4010	76287.0007	-0.462945118	-0.463370537
label_knn_norm_rfe	6000585354	57006.3000	75156.0006	-0.451834007	-0.451834007
dummy_mean_std_corr	4970105948	49667.7278	62754.428	0.603925384	0.603781128
dummy_mean_std_rfe	4860004837	48556.6167	61643.428	0.614936495	0.614936495
dummy_mean_norm_corr	5930100817	48117.1746	66286.799	-0.36293856	-0.36336398
dummy_mean_norm_rfe	5820000706	47006.0635	65175.688	-0.35182745	-0.35182745
dummy_knn_std_corr	4970126893	44186.7312	62738.095	-0.40412679	-0.404564188
dummy_knn_std_rfe	4860015782	43075.6201	61627.095	-0.40301568	-0.40301568
dummy_knn_norm_corr	5739196465	48117.4010	66287.0007	-0.362945118	-0.363370537
dummy_knn_norm_rfe	5628085354	47006.3000	65176.007	-0.351834007	-0.351834007

IV. CONCLUSION

It is noted that most of the time of a programmer for training a model is spent on data preprocessing. Various combination of data preprocessing techniques have been compared and different performance metrics have been evaluated to compare them. It can be seen that dummy coding for categorical data handling is better than label coding. Similarly, the use standardisation for feature scaling give better results most of the time. This is not any significant difference in the results after using knn or mean imputation and correlation based feature selection or recursive feature elimination.

REFERENCES

[1] Rich Caruana and Alexandru Niculescu-Mizil,"An empirical comparison of supervised learning algorithms", presented at the ICML '06: Proceedings of the 23rd international conference on Machine learning, New York, United States, Oct, 2006, pp. 161–168.

[2] S. B. Kotsiantis, D. Kanellopoulos and P. E. Pintelas, "Data Preprocessing for Supervised Learning", presented at the International journal of Computer Science, 2006, vol. 1, no. 1, pp. 111-117.

[3] Gustavo E. A. P. A. Batista and Maria Carolina Monard, 'An Analysis of Four Missing Data Treatment Methods for Supervised Learning', presented at the Applied Artificial Intelligence, Hong Kong, 2003, vol. 17, no. 5-6,pp. 519-533.

[4] Daniel J. Stekhoven and Peter Bühlmann, "MissForest—nonparametric missing value imputation for mixed-type data", presented at the Bioinformatics,Oct 28, 2012, vol. 28, no. 1, pp. 112-118.

[5] Mahsereci Karabulut Esra, Ayşe Özel Selma and Turgay brikçi, "A comparative study on the effect of feature selection on classification accuracy", presented at the Procedia Technology, 2012, vol. 1, pp. 323-327.

[6] M.A. Jayaram, Asha Gowda Karegowda and A.S. Manjunath, "Feature Subset Selection Problem using Wrapper Approach in Supervised Learning", presented at the International Journal of Computer Applications, New York, USA, 2010, vol. 1, no. 7, pp. 13-17.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:53:08 UTC from IEEE Xplore. Restrictions apply.

[7] Girish Chandrashekarand Ferat Sahin, “A survey on feature selection methods”, presented at the Computers & Electrical Engineering, New York, USA, Jan, 2014, vol. 40, no. 1, pp. 16-28.

[8] Cosmin Lazar, Jonatan Taminau, Stijn Meganck, David Steenhoff, Alain Coletta, Colin Molter, Virginie de Schaetzen, Robin Duque, Hugues Bersini and Ann Nowe, “A Survey on Filter Techniques for Feature Selection in Gene Expression Microarray Analysis”, presented at the IEEE/ACM Transactions on Computational Biology and Bioinformatics, Washington, DC, USA July, 2012, vol. 9, no. 4, pp. 1106-1119.

[9] Jose M Jerez., Ignacio Molina, Pedro J. Garcıa-Laencina, Emilio Alba, Nuria Ribelles, Miguel Martın and Leonardo Franco, “Missing data imputation using statistical and machine learning methods in a real breast cancer problem”, presented at the Artificial Intelligence in Medicine, May 12, 2010, vol. 50, pp. 105-115.

[10] Jeff Heaton, “An empirical analysis of feature engineering for predictive modeling”, presented at the Southeastcon IEEE, Norfolk, VA, USA, March 3, 2016,pp. 1-6.

[11] Pace, R. Kelley and RonaldBarry,“Sparse Spatial Autoregressions”, Statistics and Probability Letters, URL https://github.com/ageron/handson-ml, 1997

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:53:08 UTC from IEEE Xplore. Restrictions apply.

Environmental Science and Pollution Research
https://doi.org/10.1007/s11356-021-16079-4

RESEARCH ARTICLE

Emission of respirable dust from stone quarrying, potential health effects, and its management

Saurav Kumar Ambastha[1] · Anil Kumar Haritash[1]

Received: 20 May 2021 / Accepted: 17 August 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

Mining of minerals exerts adverse pressure on different compartments of environment directly or indirectly. Air is the worst affected environmental matrix, and it can carry the harmful effect of pollutants generated from mining activity even to distant places. The present study was undertaken to estimate the emission of particulate matter ($PM_{2.5}$ and PM_{10}) from different activities undertaken in stone quarrying in Mahendragarh, Haryana. The results obtained from the present study indicated that drilling, blasting, crushing, and transport of mined material are chiefly responsible for the generation of dust. Whereas drilling, blasting, and loading were responsible for emission of higher fraction of PM_{10}, crushing and re-suspension of roadside dust from movement of vehicles resulted in generation of relatively higher fraction of finer dust ($PM_{2.5}$). Modelling the transport of dust over the Hybrid Single-Particle Lagrangian Integrated Trajectory model revealed that the emitted particle may move up to the distance of about 40 km within 4 h of emission under average meteorological conditions. Fourier transform infrared (FTIR) spectroscopy analysis of dust confirmed the presence of calcite and gypsum, thus confirming the source as mining. The study concluded that generation of $PM_{2.5}$-sized particles may impose serious respiratory health effects over the workers engaged in mining, crushing, and transportation of sandstone. Apart from it, population residing downwind of the mining area is particularly vulnerable to the pulmonary effects due to inhalation of dust.

Keywords Emission · HYSPLIT model · Mining · $PM_{2.5}$ · Sandstone

Introduction

Economic growth in the Asian region has led to rapid exploitation of natural resources which has adversely affected the quality of environment, particularly around the centres of activity. To support the ongoing infrastructure development, construction of highways and railways etc., natural reserves of rock (sandstone, limestone etc.) are being indiscriminately extracted. The mines and quarries are associated with soil contamination, air pollution, and adverse effects on water resources through acid-mine drainage, leaching etc. Even the opencast quarries of non-minerals like sandstone are responsible for emission of substantial quantities of dust and other pollutants into the atmosphere resulting in negative environmental impacts over air, soil (Haritash et al. 2007), water (Chitrakshi and Haritash 2018), and vegetation (Ambastha et al. 2020). The ever-increasing demand for the materials to support the infrastructure and industrial projects has further enhanced the activity in quarries (Gulia et al. 2015). Annual average rock usage in India exceeds 1.3 billion tonnes (GIZ 2015), and it is one of the fastest growing in the world (World Bank 2017). The need for rock aggregates in India is estimated to be between 2 and 10.3 billion tonnes by the end of the next decade (GIZ 2016). Considering the environmental impacts of rock quarrying, the major effects include habitat degradation and species extinction, contamination of groundwater (Aryafar et al. 2012), emissions of dust and respirable particulate matter (Salem 2021), and other indirect mining-related environmental risks (Zobrist and Giger 2013). The site operations such as drilling, blasting, excavation, loading, unpaved road haulage, and primary crushing and stockpiling can cause substantial emission to the atmosphere. The uncontrolled emissions of fugitive dust can also cause significant environmental and health-associated concerns to the site-staff and the local community exposed to it.

Responsible Editor: Philippe Garrigues

✉ Saurav Kumar Ambastha
ambastha@gmail.com

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

Published online: 28 August 2021

The word "dust" refers to particles smaller than 75 μm in aerodynamic diameter which remain suspended in the atmosphere until it rests under its weight. Similar to other anthropogenic sources of dust emission, the mining and quarrying industries are a source of coarse as well as fine particulate matter (Hinds 1999). Natural dust (from natural resuspension and events of a sandstorm) and anthropogenic dust (from mining, crushing, etc.) are significant contributors of $PM_{2.5}$ and PM_{10} (Jain and Khare 2010), and it may be responsible for more than 50% of the accumulation of dust particularly in the mining areas. Further, the unregulated emission of dust from quarries could lead to compromised human health (Moshammer et al. 2014; Fang et al. 2019). The long-range transport of emitted dust may cause exposure and associated health effects even over the vulnerable population (Jain and Khare 2008) which is located at distant places. Studies on the environmental effects of mining of coal, minerals, and radioactive materials are many, but the availability of literature on stone quarrying and its effect on environment is very limited (Amitshreeya and Panda 2012), particularly in the Indian context. Similarly, there have been a number of reports on modelling the transport and effects of air pollution in respect of major cities (Jain and Khare 2010), but the studies on modelling the emission and transport of dust from non-mineral quarries are very limited. The present study is undertaken to determine the concentration of respirable dust (PM_{10} and $PM_{2.5}$) at locations associated with specific source of emission in a sandstone quarry and analyse the dispersion of particulate matter from quarries, which influence air quality at differing distances from the source, using the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) model.

Materials and methods

Haryana state in India is famous for agricultural production, but the location of dispersed reserves of non-minerals in the state has made stone quarrying a key activity. Six major regions Mahendragarh, Narnaul, Tosham, Bhiwani, Gurugram, and Faridabad represent areas of stone or non-mineral sand mining. Mahendragarh district is located at 28°16′N latitude and 76°9′E longitude with a geographic area of 1939 km^2 (Govt. of Haryana 2018). The climate of the region is semi-arid, represented by average maximum ambient temperature of 42°C in summers and minimum of 10°C in winters, and unevenly distributed rainfall of 500 mm during 3 months of monsoon. Mahendragarh district has nine major mining sites (Table 1), and Bakhrija stone mine is the largest with an area of about 66 km^2 (Mines and Geology Department, Haryana 2018). The location of Bakhrija mines is between 27°55′1″ to 27°54′6″ North latitude and 76°03′28.34″ to 76°03′27.56″ East longitude, and it is known for quarrying of calcite, limestone, and mica (Fig. 1).

Table 1 Mining sites located in Mahendragarh district, Haryana, India

S. no.	Place	No. of mine pit	Area (km^2)
1.	Narnaul	1	19.80
2.	Rajawas	1	53.03
3.	Amarpur Jorasi	1	10.72
4.	Bakhrija (Plot 1)	1	11.26
5.	Bakhrija (Plot 2)	1	21.65
6.	Bakhrija (Plot 3)	1	34.65
7.	Garhi	1	6.70
8.	Karota	1	9.50
9.	Mukandpur	1	3.25
Total			170.56

Source: Mines and Geology Department, Government of Haryana (2018).

Sample collection and analysis

Bakhrija quarry is surrounded by a number of stone crushing units (≈ 25 units) around it. Samples of suspended particulate matter ($PM_{2.5}$ and PM_{10}) were collected from four locations in and around Bakhrija mines during the pre-monsoon season (June 2019). During the sample collection, the ambient temperature varied from 43 to 45 °C; recorded wind speed of 0.6–1.4 m/s; humidity less than 20%; and dominant wind direction being North-West. The samples of atmospheric suspended dust ($PM_{2.5}$ and PM_{10}) were collected using the portable microprocessor-based particulate monitor (HAZDUST-EPAM 5000 model) operated for 4 h in one cycle at the flow rate of 4.0 l per min. In order to classify the variations with respect to associated mining activity, four sites, i.e., upstream of crushing units, stone crushing site, mine pit, and haul road (roadside), were identified to record dust levels due to natural background, crushing activity, mining (drilling, blasting, loading, etc.), and vehicular movement, respectively, to ascertain the relative contribution of each activity. The instrument was placed at a height of 1.5 m from ground in an open space avoiding any obstruction to the natural flow of the wind. Manual zero was performed before the start of a new series of sample collection. The values of $PM_{2.5}$ and PM_{10} were recorded at periodic intervals of 60 s for 4 h at each site, and associated activities at every site were recorded to relate the variation of SPM concentration to respective activities.

Modelling of particulate matter

For the modelling of transport, dispersion, and deposition of dust along the wind, the HYSPLIT (Hybrid Single-Particle Lagrangian Integrated Trajectory) model developed by NOAA was used. The model is used to determine the trajectory, direction, and distance of the air parcel during dispersion

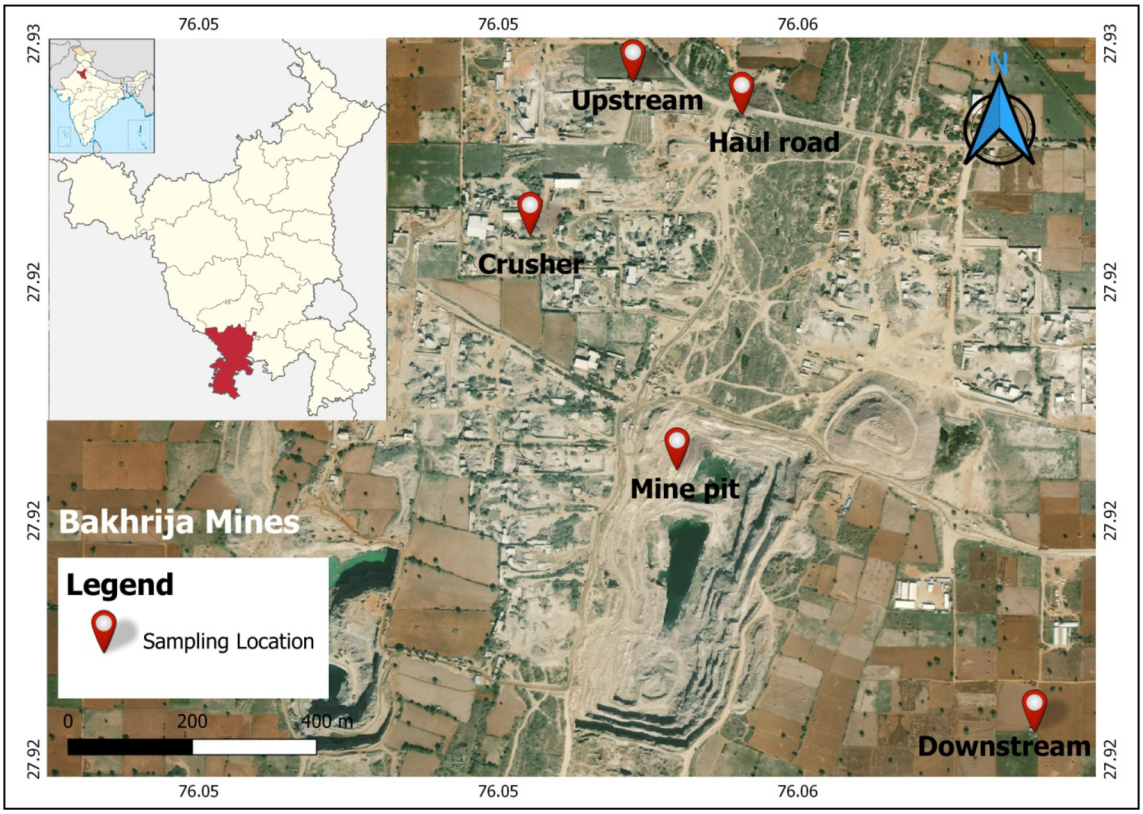


Fig. 1 Location of Bakhrija sandstone quarry, Mahendragarh, Haryana and sampling locations

along the wind with respect to time under the given set of meteorological condition. The average change in position of the particle (P_{mean}) with time ($t + \Delta t$), as per HYSPLIT, is represented as

$$P_{\text{mean}}\,(t+\Delta t) = P_{\text{mean}}\,(t)\;1/2[\,V\,(P_{\text{mean}},t) + V\,(\{\,P_{\text{mean}}\,(t) + [V\,(P_{\text{mean}},t)]\},t+\Delta t)]\,\Delta t \quad (1)$$

where "V" represents the average of the three-dimensional velocity vectors, and it is determined at the original and first-assumed positions (Draxler and Hess 1998). In this model, the particles or puffs are first released into the Lagrangian system and carried to the HYSPLIT and Global Eulerian model (GEM). The GEM and the calculations consider advection, horizontal and vertical blending, dry and wet precipitation, and radioactive decay of particles emitted (Draxler 2007).

Results and discussion

The results obtained in the present study indicate that mining and associated activities are responsible for emission of

Table 2 Concentration of the suspended particulate matter at different location of Bakhrija stone quarrying, Mahendragarh

S. no.	Location (site-specific activities)	Size	Concentration (μg/m^3)			$PM_{2.5}$/ PM_{10}
			Min.	Max.	Mean ± SD	
1	**Upstream** (Transportation only)	$PM_{2.5}$	2	6	3 ± 1.8	1.00
		PM_{10}	2	7	3 ± 1.8	
2	**Crusher** (Crushing Only)	$PM_{2.5}$	8	348	64 ± 91.1	0.71
		PM_{10}	10	484	90 ± 121.9	
3	**Mine-Pit** (Drilling, blasting, and loading)	$PM_{2.5}$	10	221	53 ± 43.9	0.52
		PM_{10}	12	224	62 ± 54.1	
4	**Haul road** (Transportation only)	$PM_{2.5}$	2	921	182 ± 215.6	0.85
		PM_{10}	6	3359	353 ± 559.2	
5	**Downstream** (No activity)	$PM_{2.5}$	2	43	8 ± 10.7	0.13
		PM_{10}	40	103	61 ± 15.0	

particulate matter which has respirable size range and can have adverse health impacts. The concentration of $PM_{2.5}$ and PM_{10} recorded at different locations at the Bakhrija stone quarry is given in Table 2.

Concentration of $PM_{2.5}$ and PM_{10}

The findings revealed that different operations in mining activity were responsible for generation of dust particles of variable size and concentration (Table 2). The concentration of $PM_{2.5}$ ranged from 02 μg/m^3 (upstream) to 921 μg/m^3 (haul road) and there was a notable variation between the concentration of $PM_{2.5}$ with reference to location and activity. The maximum concentration of $PM_{2.5}$ was recorded along the haul road (921 μg/m^3) because of the re-suspension of the dust due to movement of heavy vehicles, followed by the crushing unit (348 μg/m^3), mine pit (221 μg/m^3), and downstream the mining area (43 μg/m^3) which was located around 1.5 km downwind (Fig. 2). The concentration of PM_{10} ranged from 2 μg/m^3 to 3359 μg/m^3, and significant differences in PM_{10} concentration were observed at different locations depending on the associated activity. The maximum PM_{10} concentration of 3359 μg/m^3 was recorded along the haul road, followed by crushing unit (484 μg/m^3), mine pit (224 μg/m^3), downstream

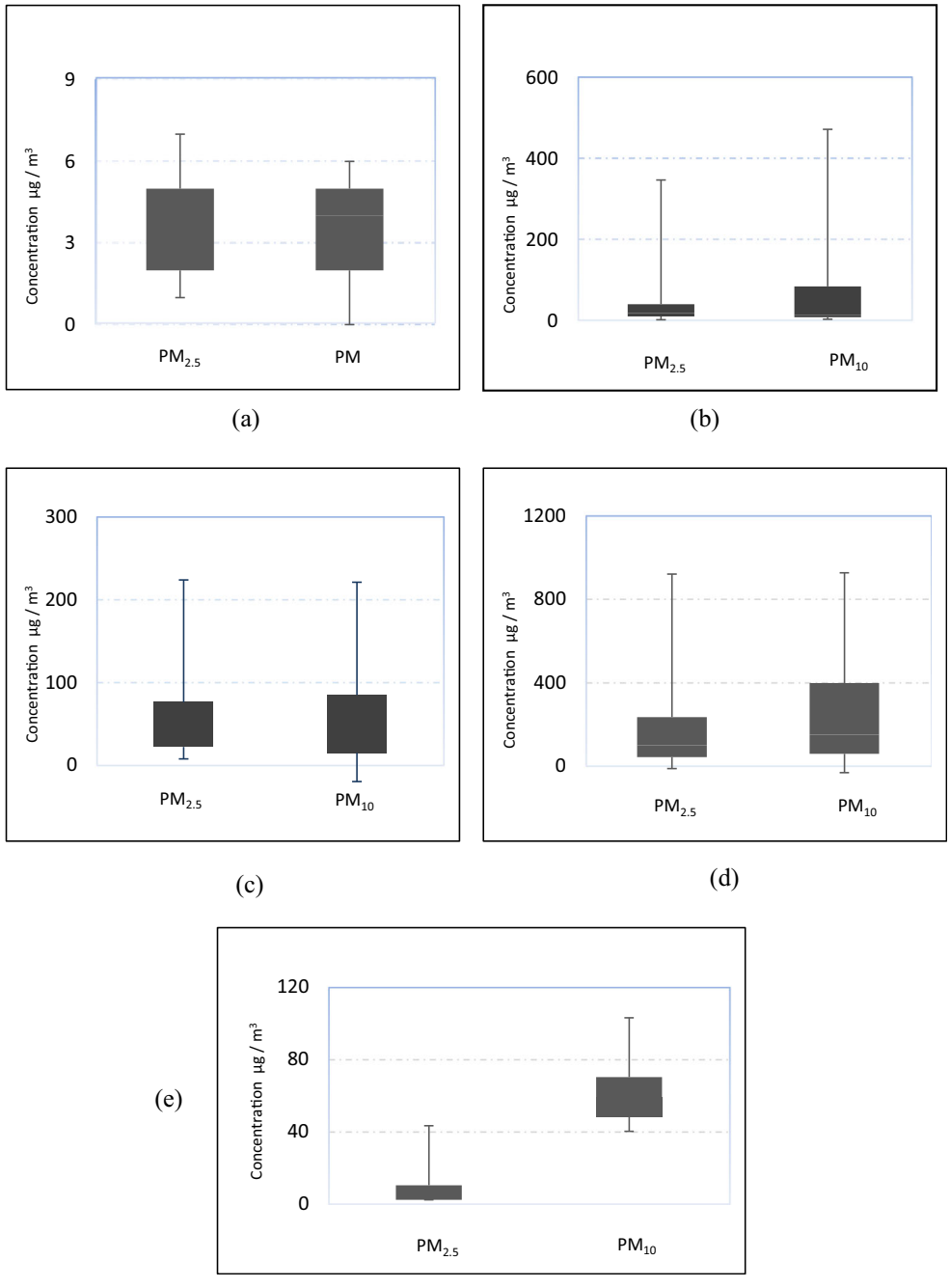


Fig. 2 Concentration of $PM_{2.5}$ and PM_{10} at (**a**) upstream of mining area, (**b**) crushing zone, (**c**) mine pit, (**d**) haul road, and (**e**) downstream of mining area

(103 μg/m^3), and upstream with level of 7 μg/m^3. The PM_{10} concentration was higher than $PM_{2.5}$ at all the locations and the SPM level significantly higher along the haul road and at the stone-crushing units, thus indicating that hauling of the mined material and its crushing are dominantly responsible for generation and emission of fine dust at stone quarrying in Mahendragarh.

There were significant variations in the relative average mass concentration of $PM_{2.5}$ and PM_{10} across the study area. The mean $PM_{2.5}/PM_{10}$ ratio values greater than 0.5 indicate that more than 50% of the respirable dust is contributed by fine particles (Das et al. 2006) (Table 2). The ratio was more than 0.5 at most places suggesting the fraction of fine particles ($PM_{2.5}$) is relatively more, except the downstream location which represents about 87% part of coarser inhalable particles compared to the finer fraction. It was observed that the hauling (resuspension) and crushing activities are dominantly responsible for emission of fine particles. Whereas the coarser particles are also emitted from different activities, their suspension and movement are restricted owing to the weight associated. On the other hand, the fine particles remain suspended and form major fraction of the suspended dust. The study confirms that hauling and stone crushing of the mined material is chiefly responsible for generation of fine dust which may impose serious health effects to the exposed workers.

Dispersion modelling of the dust

The dust generated from Bakhrija stone quarries was observed to be dominantly fine with respect to particle size and was modelled over the HYSPLIT PUFF model to determine the transportation trends, forward dispersion pattern, and distance to deposit over the ground. The meteorological data like wind speed (1.4 m/s), wind direction (North-Westerly), stability class etc. during the period of sampling was obtained from the India Meteorological Department (IMD). It was assumed that mass emission from the source is 1.0 g/h, stability class moderately unstable (category B), deposition velocity of the particle being 0.1 cm/s, and particle density of 1.0 g/cm^3 for stone dust (Tartakovsky et al. 2016). The forward dispersion of dust was simulated using horizontal resolution of 10 km up to the distance of 40 km between ground level and altitude of 100 m. The time series profile of PM2.5 concentrations during dispersion and deposition is given in Fig. 3, and it is observed that the dispersion goes to a maximum distance of about 40 km before complete deposition under the observed set of meteorological conditions. Similarly, the vertical dispersion of PM particles in the atmosphere (Fig. 4) was observed to be concentrated up to a height of 4000 m during the forward dispersion. The computed results indicated that the fine fraction can travel a significant distance downstream before getting deposited, thus, exposing the population located at farther distance as well.

Mineralogical composition of dust, possible environmental effects, and management

Settling of the emitted dust over adjoining agricultural fields may induce physico-chemical alteration in the properties of soil, thereby resulting in altered pH of soil, reduced porosity, reduction in total organic carbon (TOC), and adversely affecting the cation exchange capacity (Haritash et al. 2007). The Fourier transform infrared (FTIR) spectroscopy of dust collected from the study area revealed the presence of gypsum ($CaSO_4.2H_2O$), calcite ($CaCO_3$), and kaolinite ($Al_2Si_2O_5(OH)_4$) minerals (Fig. 5), thereby ruling out the possibility of addition of toxic impurities to soil through the dustfall. Moreover, addition of 'Ca' and 'Mg' to soil would significantly improve cation exchange capacity resulting in improved productivity of such soils. Simultaneously, addition

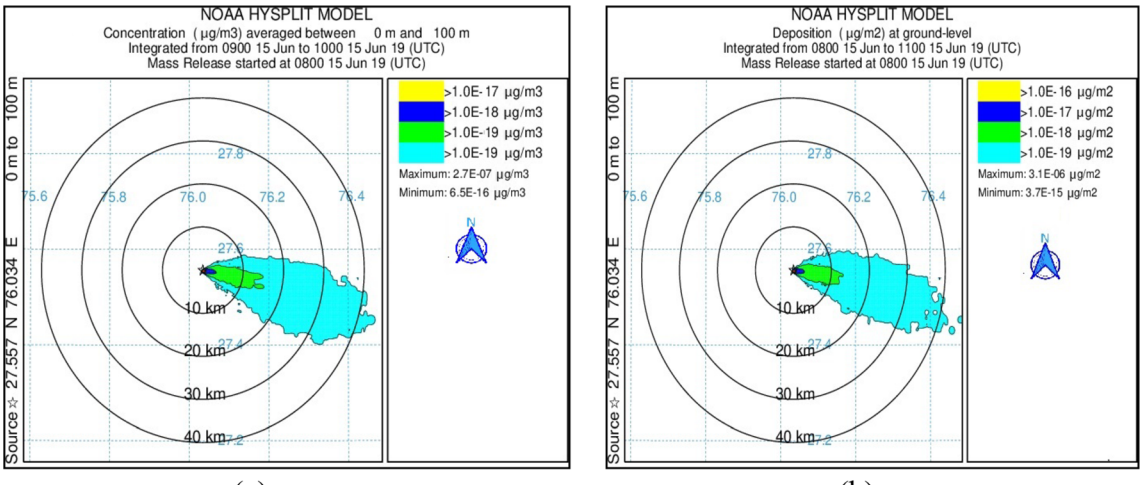


Fig. 3 (**a**) Area under dust concentration due to the source. (**b**) Area under dust deposition

of the fine dust to soil (dominantly aeolian sand) results in improved water holding capacity and maintaining the neutral pH of better crop yield in case of sandy soil being observed in the region around Bakhrija quarry. The study suggests that stone quarrying and crushing may not exert any significant adverse effect on the quality of agricultural soil around the Bakhrija stone quarries. Considering the fine ($PM_{2.5}$) nature of emitted dust, the health effects over the exposed population could be cardiac or lung disease, non-fatal heart attacks, erratic heartbeat, asthma, reduced lung capacity, elevated respiratory symptoms like airway inflammation, coughing, or trouble breathing (Jain and Khare 2008). So, there may be a possibility that people working in Bakhrija quarry and stone crusher units may suffer from these occupational issues.

The effects which are likely to be caused by the mining activities are to be minimised through an effective management plan. Suitable measures should be taken up to regulate the emissions of dust during material handling, crushing, and transportation. During transport, the material should be covered with tarpaulin, and water sprinkling should be operated on haul roads to reduce fugitive emissions. The dispersion of dust to distant places can be restricted by developing a green

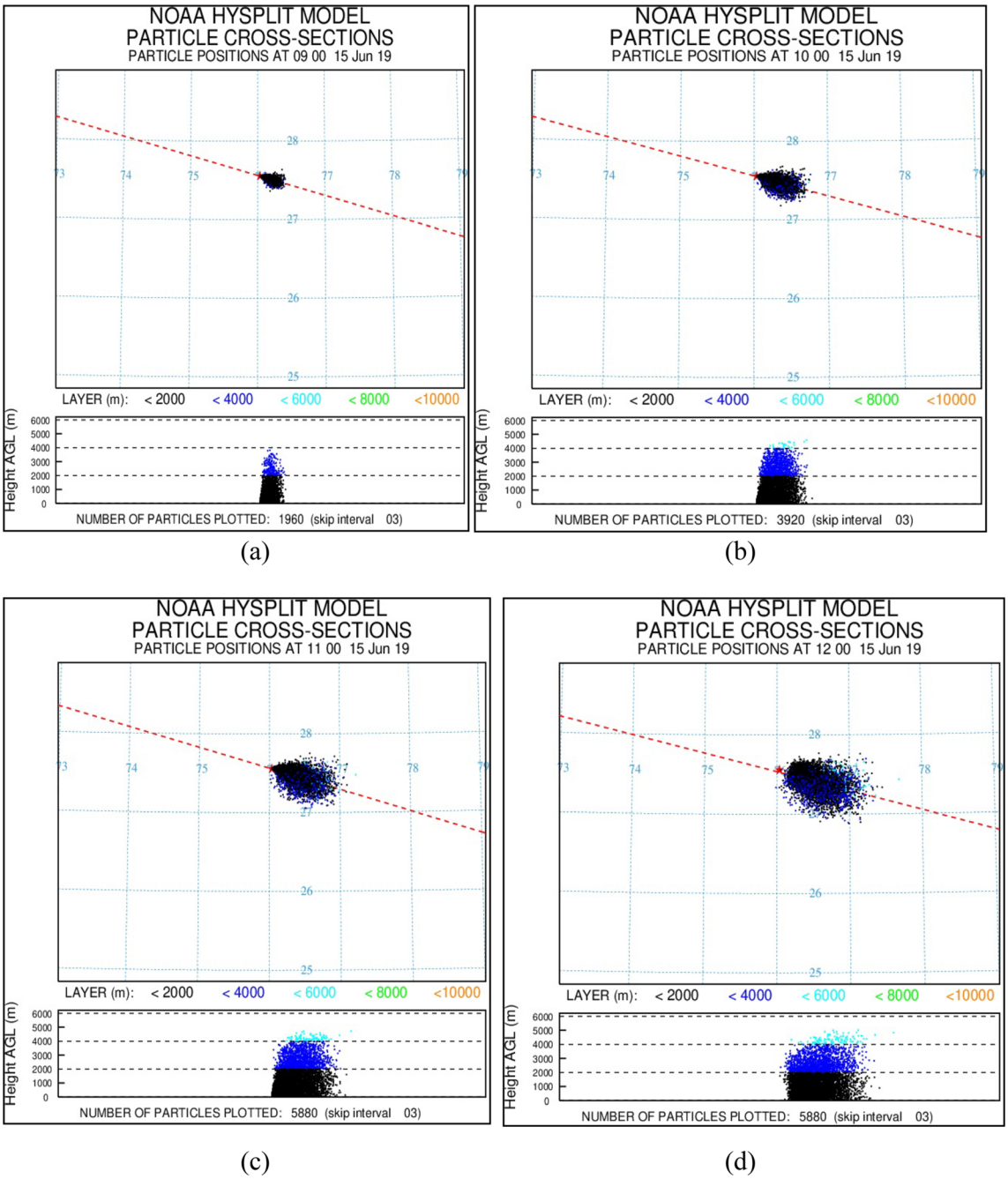


Fig. 4 Particle position with altitude and deposition with time (**a**–**d**)

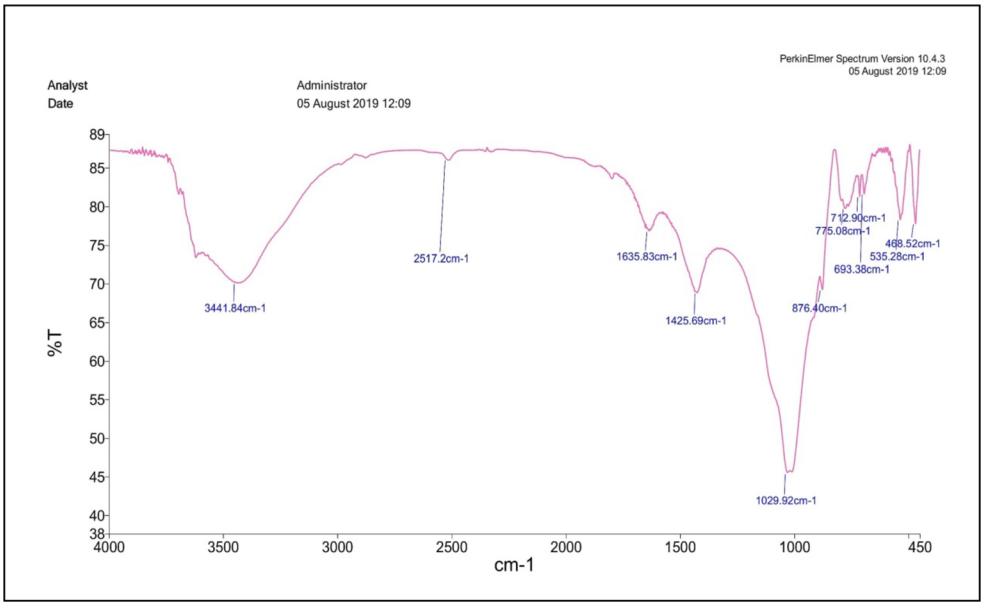


Fig. 5 FTIR spectrum of mine dust representing the presence of calcite (712, 876, 1425, 2517) and gypsum (1635, 3441) minerals

belt around the mining and crushing zone using the vegetation local to the area. It will substantially help preserve the regional ecological balance, soil and hydrological conditions, and air quality as well. On the occupational front, vocational training is to be given to workers, and safety equipment should be provided. The health of workers during mining shall be ensured in compliance with the regulations.

Conclusion

Based on the result obtained in the present study, it is concluded that hauling of mined material and crushing are associated with generation of fine dust particles ($PM_{2.5}$) which may travel long distance in the direction of wind resulting in exposure of vulnerable population even at distant places. Although the dust fall does not cause any significant adverse effect on soil owing to its chemical composition, its fine particle size has got deeper penetration capability in the respiratory system if humans are exposed. Workers engaged in crushing, drilling, crushing, and loading crushed material are most exposed. Simultaneously, population present downwind villages may be affected if effective steps to ameliorate the dust dispersion are not taken up. Establishing a thick green belt comprising indigenous resistant trees may be an eco-friendly option to maintain the air quality around such stone-quarries and crushing units.

Acknowledgements The authors are thankful to Mr. Rajesh Sehrawat, Mine Inspector, Mahendragarh for his help in several ways during this study.

Author contribution SKA: conceptualisation, methodology, software, data curation, writing- original draft preparation, software, validation, writing- reviewing and editing. AKH: conceptualisation, methodology, visualisation, investigation, supervision, original draft preparation, validation, writing- reviewing and editing.

Data availability The authors declare that (the/all other) data supporting the findings of this study are available within the article (and its supplementary information files). In case, the reviewers require any data which support the findings of this study are available on request from the corresponding author.

Declarations

Ethics approval and consent to participate Authors declare no involvement of human and animals in the research work.

Consent for publication The authors give their consent for the publication of the manuscript, which can include photograph(s) and/or videos and/or case history and/or details within the text ("Material") to be published in the Environmental Science and Pollution research Journal.

Conflict of interest The authors declare no competing interests.

References

Ambastha S, Negi S, Singh A, Ghunasham M, Haritash AK (2020) Evaluation of Air Pollution Tolerance Index and Anticipated Performance Index of trees in and around the stone quarrying region of Mahendragarh, Haryana. Indian J Air Pollut Control 20(1-2):17–27

Amitshreeya R, Panda RB (2012) Dust pollution in stone crusher units in and around Balasore, Orissa, India. J Ind Pollut Control 28(1):41–44

Aryafar A, Yousefi S, Doulati AF (2012) The weight of interaction of mining activities: groundwater in environmental impact assessment

using fuzzy analytical hierarchy process (FAHP). Environ Earth Sci 68:2313–2324. https://doi.org/10.1007/s12665-012-1910-x

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

Das M, Maiti SK, Mukhopadhyay U (2006) Distribution of $PM_{(2.5)}$ and $PM_{(10\text{-}2.5)}$ in $PM_{(10)}$ fraction in ambient air due to vehicular pollution in Kolkata megacity. Environ Monit Assess 122(1-3):111–123. https://doi.org/10.1007/s10661-005-9168-3

Draxler RR (2007) Demonstration of a global modeling methodology to determine the relative importance of local and long-distance sources. Atmos Environ 776–778. https://doi.org/10.1016/j.atmosenv.2006.08.052

Draxler RR, Hess GD (1998) An overview of the HYSPLIT_4 modeling system for trajectories, dispersion, and deposition. Aust Meteor Mag 295–308

Fang D, Huang W, Antkiewicz DS (2019) Chemical composition and health risk indices associated with size-resolved particulate matter in Pearl River Delta (PRD) region. China Environ Sci Pollut Res 26:12435–12445. https://doi.org/10.1007/s11356-019-04618-z

GIZ (2015) Progress report on sustainability. Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, Germany

GIZ (2016) Climate and Environmental Report. Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, Germany

Government of Haryana (2018) https://mahendragarh.gov.in/about-district/. Accessed 26 Aug 2019

Gulia S, Khare M, Nagendra Shiva SM, Khanna I (2015) Urban air quality management—a review. Atmos Pollut Res 6(2):286–304. https://doi.org/10.5094/APR.2015.033

Haritash AK, Baskar R, Sharma N, Paliwal S (2007) Impact of slate quarrying on soil properties in semi-arid Mahendragarh in India. Environ Goel 51:1439–1445. https://doi.org/10.1007/s00254-006-0440-9

Hinds WC (1999) Atmospheric aerosols. In: Hinds WC (ed) Aerosol technology: properties, behaviour, and measurement of airborne particles, 2nd edn. Wiley- Interscience, USA, pp 304–315

Jain S, Khare M (2008) Urban air quality in mega cities: a case study of Delhi City using vulnerability analysis. Environ Monit Assess 136:257–265. https://doi.org/10.1007/s10661-007-9681-7

Jain S, Khare M (2010) Adaptive neuro-fuzzy modelling for prediction of ambient CO concentration at urban intersections and roadways. Air Qual Atmos Health 3:203–212. https://doi.org/10.1007/s11869-010-0073-8

Mines and Geology department Government of Haryana (2018) http://minesharyana.gov.in/en/production-report-mahendergarh-0

Moshammer H, Hochgatterer K, Haluza D (2014) Occupational exposure to mineral dust. Effects of on lung function in a nine-year study. Occup Environ Med 71:A62. https://doi.org/10.1136/oemed-2014-102362.194

Salem HS (2021) Evaluation of the stone and marble industry in Palestine: environmental, geological, health, socioeconomic, cultural, and legal perspectives, in view of sustainable development. Environ Sci Pollut Res 28:28058–28080. https://doi.org/10.1007/s11356-021-12526-4

Tartakovsky D, Stern E, Broday DM (2016) Dispersion of TSP and PM(10) emissions from quarries in complex terrain. Science Total Environment 15;542(Pt A):946–54. https://doi.org/10.1016/j.scitotenv.2015.10.133

The World Bank Annual Report (2017). https://www.worldbank.org/en/about/annual-report. Accessed 26 August 2019

Zobrist J, Giger W (2013) Mining and the environment. Environ Sci Pollut Res 20:7487–7489. https://doi.org/10.1007/s11356-013-1990-5

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

Environmental Science and Pollution Research
https://doi.org/10.1007/s11356-021-15602-x

GREEN ENERGY FOR ENVIRONMENTAL SUSTAINABILITY

Emissions from homogeneous charge compression ignition (HCCI) engine using different fuels: a review

Sanjeev Kumar Verma [1] · Subhashish Gaur [1] · Tabish Akram [1] · Samsher Gautam [2] · Anil Kumar [1,3]

Received: 11 April 2021 / Accepted: 19 July 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

The main objective of this paper is to review the emission characteristics of homogeneous charge compression ignition (HCCI) engines using different fuels and additives. Additionally, the impacts of various operating conditions on HCCI engine emissions are also analyzed. There is a demand faced by the engine and car manufacturers for high fuel efficiency and low emission from both consumers and the government. HCCI is an alternative internal combustion technology that is more efficient and cleaner than traditional combustion techniques. In this engine, fuel and oxidiser (typically air) are entirely mixed and compressed to the point of auto-ignition. It is a new combustion concept that provides high efficiency and resolves high nitrogen oxide (NOx) and particulate matter (PM) emissions simultaneously. Conventional diesel engine emits NOx around 2000–2500 ppm which is a primary concern; however, when diesel is used in HCCI engine, it reduces NOx to only 70–400 ppm. Furthermore, natural gas cannot be used as fuel in the conventional engine but can be used in HCCI engine. It helps reduce NOx emission even below 10 ppm. This paper reviews several works done by researchers on HCCI engines to reduce their emissions and benefit future research.

Keywords HCCI engine · Fuel · Additive · Emission · Pollution · Environment

Introduction

Spark ignition (SI) and compression ignition (CI) engines are the most widely used ground transportation engines these days. Even though both engines are reciprocating in nature, they are still different in fuel economy, power efficiency, and emissions. SI (petrol) and CI (diesel) engines use fossil fuels and contribute largely to air pollution in urban areas. A large quantity of NOx is formed in the conventional type of IC engines because of high temperature inside the cylinder. Smoke is formed within the combustion chamber due to localized fuel-rich regions and localized low-temperature zone. Photochemical smog is formed due to this NOx's reaction with atmospheric gases (Hussaini et al. 2016). PM emitted from diesel engines causes various respiratory problems like asthma, chronic obstructive pulmonary disease (COPD), lung cancer, and many more. Many emission standards are being implemented due to such adverse effects on human health because of these pollutants. They require the reduction of pollutants like PM and NOx emissions (Ohur and Kariuki 2014).

A new combustion system is required to minimize the emission level from conventional internal combustion (IC) engines. HCCI is a technique that is becoming more popular because of its capability of reducing PM and NOx emissions simultaneously by using compression ignition of the homogeneous fuel-air mixture. It is a form of IC engine in which oxidizer (mainly air) is thoroughly mixed with fuel and compressed to the point of auto-ignition. Energy released from this exothermic reaction is transformed in an engine in work and heat like other combustions. HCCI combines characteristics of conventional petrol engines and diesel engines (Janathanan et al. 2016).

HCCI engine gives improved thermal efficiency and very less emissions. It can be implemented by modifying either

Responsible Editor: Philippe Garrigues

✉ Anil Kumar
anilkumar76@dtu.ac.in

[1] Department of Mechanical, Production & Industrial Engineering and Automobile Engineering, Delhi Technological University, -110042, Delhi, India

[2] Harcourt Butler Technical University, -208 002, Kanpur, India

[3] Centre for Energy and Environment, Delhi Technological University, -110042, Delhi, India

Published online: 03 August 2021

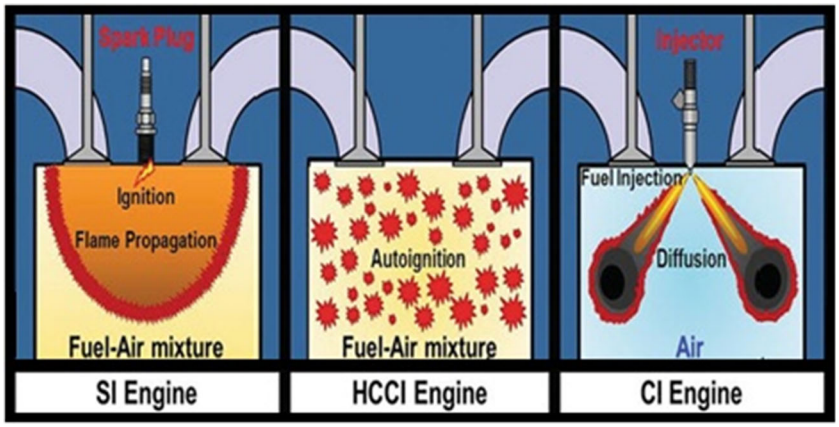


Fig 1 Comparison of SI, HCCI, and CI engines

spark ignition or compression ignition engine using any fuel type for combination. Air-fuel mixture is kept lean in HCCI engines. It automatically ignites in different locations and burns volumetrically without recognizable flame propagation. In HCCI engine, fuel delivery is essential for the controlled combustion process.

HCCI engine is a four-stroke IC engine that works similarly to SI and CI engines. Fuel is injected into each cylinder, and air is introduced into the cylinder through a separate air intake manifold during suction stroke. Fuel and air are mixed perfectly. Temperature starts increasing inside the combustion chamber due to compressed air-fuel mixture during the compression stroke. As sufficient heat is accumulated for combustion, ignition takes place without any use of a spark plug. The combustion in an HCCI engine is a lean and low temperature process (unlike SI and CI engines), and its energy is released across the entire combustion chamber. Complete air-fuel mixture is burned and produces a similar amount of power (as SI or CI engine) by using lesser amount of fuel and releasing very few emissions. At the end of the power stroke, an exhaust stroke is initiated. Still, exhaust valves close before all of the burned gases can be evacuated, trapping some amount of latent heat of combustion. This heat is preserved, and small amount of fuel is injected inside the combustion chamber through the inlet valve for pre-charging before the next intake stroke begins (Mulane and Limaye 2016) (Fig. 1).

Lean combustion in HCCI increases the fuel efficiency by 15% over the conventional SI engine. It has lower emissions (especially NOx) than a conventional CI engine. Fuel is burned quicker and at a lower temperature which reduces heat energy losses. Also, the overall engine noise, vibration, and harshness are lower than conventional engines. HCCI engine does not require any external aid to initiate the combustion process. A variety of fuels can be used in the HCCI engine, which also helps in reducing emission losses. Low-pressure fuel injection system is used, and a spark plug is also not required, making manufacturing cheaper. Also, it has the least throttle losses (Hasan and Rahman 2016).

Although HCCI has an advantage over conventional engines in terms of thermal efficiency and NOx emission, its combustion has many difficulties. Combustion in HCCI engines is achieved by controlling the pressure, temperature, and composition of the air-fuel mixture to ignite itself when the piston reaches top dead centre (TDC). It is more complex than a direct control mechanism like a fuel injector in the CI engine or a spark plug in the SI engine, which controls ignition timing. High cylinder pressure in HCCI needs more robust engine construction, which makes it expensive. Otherwise, there is a risk of explosion due to high energy produced during power stroke. Also, it has limited power and a small torque range compared to SI engine. Although it has less NOx emission, high hydro-carbon (HC) and carbon monoxide (CO) emissions are compared to the SI engine. HCCI engine has weak cold start capability, and difficulty is also faced in preparing homogeneous charge mixture (Venkataramana 2013).

Conventional fuels like petrol and diesel are the main causes of emissions from automobiles. It is understood that HCCI technology offers better fuel economy and emissions control; however, it is not confident that these engines will deliver these characteristics inexpensively. Continued advancement is necessary to bring it into the daily production of the vehicle. These challenges must be overcome for the mass production and implementation of HCCI engines in the daily production of the vehicle. This paper reviews several works done by researchers on the HCCI engine to reduce its emission. Prospective fuels for HCCI engines to get minimum emission are also discussed. This compilation will be beneficial for carrying out future research in this field.

Emissions from HCCI engine

Natural gas

Yao et al. (2006) performed a study to find controlling strategies of HCCI combustion fuelled with methanol and

dimethyl ether. Experiments were performed on a single-cylinder–modified direct-injection diesel engine which was water-cooled. NOx emissions were observed very low for the HCCI engine at 1400 rpm, and 0.64 MPa indicated mean effective pressure (IMEP) and varying die-methyl ether (DME) and exhaust gas recirculation (EGR). Till 10% EGR, NOx emission was around 30 ppm. At higher EGR like 25%, 40%, and 65%, NO_X emission was very less, even less than 10 ppm. NOx emission increases sharply when knocking occurs up to 70 ppm. HC and CO emissions in HCCI engine were analyzed at 1400 rpm and 0.44 MPa IMEP and varying DME and EGR. On increasing EGR rate from 0 to 65%, keeping DME at around 62%, HC and CO emission was around 0.3–0.4%. DME was increased to around 74%; HC emissions reduced to less than 0.15%, and CO emissions reduced to less than 0.1%. Exhaust gas recirculation and dimethyl ether have a considerable effect on the performance and emission of HCCI engine.

He et al. (2013) studied the combustion characteristics and emission characteristics of an HCCI engine fuelled by n-butanol. Experiments were performed on a four-stroke single-cylinder gasoline engine equipped with variable valve lift (VVL) and variable valve timing (VVT) devices on both the valve intake and exhaust. It was observed that for 1500 rpm, 63° CA after top dead centre (ATDC), and 0.26 MPa IMEP, NOx emission was 40–60 ppm whereas for 2000 rpm, 43° CA ATDC and 0.158 MPa IMEP, it was 10–20 ppm. Very less amount of NO emission was observed as compared to other conventional engines. Methanol emission was around 15–16 ppm for 1500 rpm and 13 ppm for 2000 rpm; however, ethanol emission was 345–360 ppm for 1500 rpm and 270–300 ppm for 2000 rpm.

Maurya and Agarwal (2010) performed experiments to study the combustion and emission characteristics of HCCI combustion engine running on ethanol. A four-stroke two-cylinder engine was used, which was naturally aspirated, air-cooled, and bowl-shaped combustion chamber. Direct injection diesel engine was modified for experiments. An experiment was performed at 1500 rpm, and intake air temperature of fuel was maintained between 120 and 150°C to analyze different emission parameters. HC emission was 200–400 ppm at 120 °C for a range of "λ" value from 2 to 3.5. However, at 150 °C, it was 175–225 ppm for a range of "λ" value from 3 to 5. CO emission ranges from 0.1 to 0.8% for different temperatures and different "λ" values. NOx emission for all the temperatures was found to be less than 10 ppm. Results show that NOx emissions are lower than 10 ppm for all stable operation points; however, HC and CO emissions were higher.

Tsolakis et al. (2007) studied the exhaust gas fuel application of reforming in diesel and HCCI engines fuelled with biofuels. They explained the exhaust gas fuel-reforming benefits by adding simulated reformed gas to a diesel engine which was fuelled by the mixture of 50% ultra-low sulfur diesel and 50% rapeseed methyl ester, referred to as B50. They also used a HCCI engine that uses bio-ethanol. It was observed that at 4.5 bar IMEP and 2.6 "λ", exhaust gas composition was 0.01% CO, 12.6% O_2, 4.76% CO_2, and 5.02% H_2O. Also, as the percentage by volume of EGR increases from 0 to 20%, NOx emission is reduced from 700 to 400 ppm, and on increasing reformed EGR (REGR) by same, NOx emission reduced from 700 to 650 ppm for fuel injection at standard 22° BTDC. As fuel injection is retarded to 19° BTDC, NOx emission is found to be 700–550 ppm. At 6.1 bar IMEP and 2.14 λ, exhaust gas composition was 0.01% CO, 10.8% O_2, 6.3% CO_2, and 6.4% H_2O. NOx emission decreased from 1050 to 600 ppm for EGR and 1050 to 800 ppm for REGR.

Olsson et al. (2001) worked on a dual-fuel turbocharged HCCI engine. A 6-cylinder truck engine was modified for dual fuel turbocharged HCCI engine operation. Ignition timing was controlled by using two different fuels, ethanol and n-heptane. A study was done to demonstrate high load operation of a HCCI engine. As the load was increased at high speed, HC emission reduced from 10 to 8 g/kWh, and at low speed, it was around 6–7 g/kWh. The same trend can be seen in CO emission also. At high speed, CO emission reduced from 90 to 20 g/kWh on increasing the load; and at low speed, CO emission was around 10–15 g/kWh. It was observed that NOx emission increases with an increase in load; however, even at the highest load condition of 16 bar break mean effective pressure (BMEP), NOx was less than 4 ppm (0.063 g/kWh).

Again Olsson et al. (2003) performed experiments to study the cooled EGR effect on emissions and performance of turbocharged HCCI engine. A six-cylinder D12 truck engine was modified for this study. Ethanol and n-heptane port fuel injection was used. For the closed-loop combustion control, cylinder pressure sensors were also used. It was observed from the experiment that, for intermediate loading of 11.5 bar, HC emission reduces from 40 to 20g/kg_{fuel} when EGR is increased from 0 to 60%; and for intermediate loading of 7.5 bar, HC emission reduces from 25 to 10 g/kg_{fuel}. Maximum CO emission is 22 g/kg_{fuel} for both 11.5 and 7.5 bar load conditions and reduces to around 10 g/kg_{fuel} on increasing EGR rate. Observing for NOx emission at 11.5 bar, it was found that it ranges between 0.02 and 0.07 g/kWh. When a turbocharger is used at 1600 rpm and 11 bar BMEP, NOx emission for no EGR condition is between 15 and 25 ppm and for 45% EGR, NOx ranges between 20 and 35 ppm.

Polat (2015) performed experiments to study engine performance, combustion, and emission characteristics of HCCI engine using diethyl ether blend and ethanol. Four-stroke single-cylinder HCCI engine with port injection was used. Mixture of diethyl ether and ethanol was used in the experiment in the ratio like 30% ethanol–70% diethyl ether (E30/D70), 40% ethanol–60% diethyl ether (E40/D60), 50%

ethanol–50% diethyl ether (E50/D50), and 100% diethyl ether (DEE). Observations were taken at a fixed speed of 1200 rpm. For E30/D70 mixture, if inlet temperature was fixed at 353 K and "λ" is increased from 1.5 to 3, CO emission reduced from 0.09 to 0.07% and HC emission reduced from 230 to 180 ppm. However, on fixing "λ" at 2 and increasing temperature from 333 to 393 K, CO emission reduced from 0.05 to 0.07%, and HC emission reduced from 330 to 240 ppm. For E40/D60 mixture, on increasing λ from 1.46 to 2.25 at same 353 K, CO emission reduced from 0.08 to 0.05%, and HC emission reduced from 290 to 220 ppm. But on fixing "λ" at 2 and increasing temperature from 333 to 393 K, CO emission reduced from 0.14 to 0.1%, and HC emission reduced from 425 to 360 ppm. For E50/D50 mixture, if inlet temperature was fixed at 353 K and "λ" is increased from 0.8 to 1.6, CO emission reduced from 0.15 to 0.11%, HC emission reduced from 340 to 300 ppm, and NOx emission ranges between 94 and 40 ppm. But on fixing "λ" at 1 and increasing temperature from 333 to 373 K, CO emission reduced from 0.22 to 0.16%, HC emission reduced from 460 to 400 ppm, and NOx emission increased 75 to 117 ppm. For DEE, if inlet temperature is fixed at 353K and increased "λ" from 1.59 to 2.85, CO emission reduced from 0.065 to 0.04% and HC emission reduced from 190 to 140 ppm.

Ying et al. (2009) performed a comparative study of emission and combustion characteristics of HCCI engine and HCCI-DI engine. A four-stroke single-cylinder naturally aspirated direct injection diesel engine fuelled with dimethyl ether was used for the experiment. Observations were taken for different values of "*r*" which was ratio of the mass consumption rate of port aspirated DME fuel and total mass of fuel. For 18% "*r*" ratio, NO_X emission increases from 150 to 350 ppm on increasing BMEP, and as "*r*" ratio is increased to 1, NOx emission decreases to almost 10 ppm. Maximum NOx emission was found to be 700 ppm for direct injection compression ignition (DICI) engine. Maximum HC emission reduced from 1080 for 100% "*r*" ratio to 300 ppm for 18% "*r*" ratio at low load condition. As load is increased, HC emission was further reduced to 500 ppm for 100% "*r"* ratio and almost 0 ppm for other "*r*" ratio values. On increasing load from 0.05 to 0.35 MPa BMEP, CO emission reduced from 2.2 to 0.1% for 100% "*r*" ratio. But for 18%, 26% and 33% "*r*" ratio, CO emission increases on increasing the load. As EGR was increased from 0 to 12%, NOx emission reduced from 50 to 20 ppm for 33% "*r*" value and from 72 to 60 ppm for 26% "*r*" value. HC emission increased from 350 to 450 ppm for 26% "*r*" value and from 440 to 520 ppm for 33% "*r*" value. CO emission increased from 0.2 to 0.35% for 26% "*r*" value and from 0.45 to 0.6% for 33% "|*r*" value.

Yap et al. (2005) experimented with studying exhaust gas fuel reforming on HCCI engine operation worked on natural gas. Modified single-cylinder SI engine was used for compression ratio from 10.5 to 15 at 1500-rpm engine speed. CO emission increased with REGR addition was observed during emission characteristics. CO emission is 0.16% for 10% H_2 in 5% REGR, 0.18% for 15% H_2 in 5% REGR, and 0.06% for HCCI combustion without H_2 supply. Similarly, HC emission also increased with the addition of H_2 in REGR. HC emission increased to 210 ppm if 15% H_2 was supplied with 5% REGR and 200ppm at 10% H_2 with 5% REGR. But NO_X emission decreases on supplying H_2 with REGH. NO_X emission reduced to 300 ppm if 15% H_2 was supplied with 5% REGR and 500 ppm at 10% H_2 with 5% REGR compared to 800 ppm for no EGR condition.

Maurya and Agarwal (2014) modified a four-cylinder four-stroke engine and experimented with different engine speeds by preparing homogeneous charges using port fuel injection. Air-fuel mixture was auto ignited by air preheater intake in the combustion chamber. Heat release curve for calculating combustion parameters was used. NOx emission was below 10 ppm till 1800°C of mean cylinder temperature, but after that, it boosts to 70–90 ppm; however, in conventional mode, NOx emission ranges around 2000–2500 rpm. Highest CO emission was observed to be 55 g/kWh for methanol. However, it was lower in the case of gasoline-fuelled engine. It was found that HC emission was reduced by advancing the combustion timing. The highest HC emission was 18.4 g/kWh at the lowest engine load for gasoline in HCCI mode.

Mack et al. (2009) investigated heat release rates, intake temperatures, engine operating limits, and exhaust emission parameters for the engine operating with 40%, 60%, 80%, 90%, and 100% ethanol-in-water mixtures. It was observed that HC emission was above 1000 ppm for E60/W40 mixture, between 400 and 600 ppm for E80/W20 mixture, and less than 200 ppm for 100% ethanol fuel. Similarly, CO emission was also highest for E60/W40 mixture going maximum to 1400 ppm and least for 100% ethanol fuel around 500 ppm. Also, NOx emission for 100% ethanol fuel was around 0.5–4 ppm; and for E60/W40 mixture, it was 4–8 ppm. It was concluded that CO, HC, and NO emissions increased, but O_2 emission decreased with increasing water content in fuel.

Gharehghani (2018) investigated three fuels: natural gas, ethanol, and methanol, for finding load limits vis-a-vis proper operating range for each type of fuel. A single-cylinder variable compression ratio diesel engine of 507cc volumetric capacity was used for experiments. In this study, for stable HCCI operation, a high load limit was defined by the knock phenomenon and low load limit by combustion stability. CO emission for natural gas ranges between 1.5 and 6.5 g/kWh, whereas for ethanol and methanol, CO emission ranges between 0.7 and 5.8 g/kWh, which was less than natural gas. Similarly, HC emission for natural gas was around 0.8–4.2 g/kWh, whereas for ethanol and methanol, it ranges between 0.35 and 4.9 g/kWh. NO_X emission for natural gas was around 0.02–0.12 g/kWh, whereas for ethanol and methanol, it was around 0.06–0.26 g/kWh and 0.05–0.45 g/kWh, respectively.

Therefore, ethanol and methanol could be used for the lean mixture (low loads) and lower intake temperatures. In contrast, natural gas should be used for HCCI for the rich mixture (high loads) and higher intake temperatures. Methanol is also usable even for ultra-lean mixtures. Emissions from HCCI for Natural Gas are mentioned Table 1.

Diesel

Singh and Agarwal (2012) performed an investigational experiment on emission, combustion, and the HCCI engine performance worked on diesel. Experiment was performed on an engine of two cylinders in which one cylinder was modified to operate on HCCI technique. In contrast, other cylinder was operated on conventional CI engine mode. Combustion, emission, and performance characteristics were analyzed with 15% EGR and without EGR. As data of emissions are checked, it was found that maximum CO emission in diesel engine is 5 g/kWh, whereas for HCCI engine, it is 7 g/kWh; and on using 15% EGR in HCCI engine, CO emission further increases to 13 g/kWh. Also, the maximum HC emission in diesel engine is 0.75 g/kWh, whereas for HCCI engines, it was 2.1 g/kWh; and on using EGR in HCCI engine, HC emission further increases to 3 g/kWh. On the contrary, NOx emission in diesel engine was 8.9 g/kWh and for HCCI engine, it decreased to 3.9 g/kWh; and if 15% EGR was used in HCCI engine, NOx emission further decreases to 2.9 g/kWh. CO and HC emission increases in HCCI engine, whereas NOx emission decreases compared to diesel engine emissions.

Gowthaman and Sathiyagnanam (2016) experimented with analyzing the various effects of inlet air temperature and fuel injection pressure on performance and emission behavior of HCCI engine worked on diesel. Inlet air temperature was varied from 40 to 70°C, and injection pressure in the port fuel injector is varied from 4 to 5 bar. It was observed that at 4 bar, the maximum NOx emission of the conventional diesel engine is 980 ppm, and for HCCI engine worked on diesel, it was 520 ppm at 60 °C inlet air temperature. For higher pressure fuel injection of 5 bar, NOx emission for diesel remained the same, but for HCCI engine, it increased to 650 ppm. Also, if BMEP was increased, NOx emission increases. If inlet temperature of HCCI engine was increased by 40 °C, 50 °C, and 60 °C, a rise in NOx emission can be seen by 2%, 5%, and 9%, respectively. HCCI engine, operating at 4-bar injection pressure and 50 °C inlet temperature, emits smoke density maximum of 32 HSU (Hatridge Smoke Unit), whereas conventional diesel engine emits smoke density ranging between 36 and 73 HSU. If inlet temperature of HCCI engine is increased by 40 °C, 50 °C, and 60 °C, rise in smoke emissions can be seen by 2%, 4%, and 5%, respectively. HC emission in the case of the diesel engine is 50 ppm and HCCI engine was around 100 ppm. CO emission in diesel engines is 0.2%, and HCCI engine was around 0.34–37%. As the inlet temperature increases from 40 to 60 °C, HC and CO emissions reduce at all load conditions.

Table 1 Emission values for natural gas

Emission parameters Authors	NO_X	HC	CO	CO_2	O_2	H_2O
Yao et al. (2006)	0–70 ppm	0.15–0.3%	0.1–0.4%	-	-	-
He et al. (2013)	10–60 ppm	13–360 ppm	-	-	-	-
Maurya and Agarwal (2010)	0–10 ppm	175–400 ppm	0.1–0.8%	-	-	-
Tsolakis et al. (2007)	550–1050 ppm	-	0.1%	4.76–6.3%	12.6–10.8%	5.02–6.4%
Olsson et al. (2001)	0–0.63 g/kWh	6–10 g/kWh	10–90 g/kWh	-	-	-
Olsson et al. (2003)	15–35 ppm	10–40 g/kg	10–22 g/kg	-	-	-
Polat (2015)	40–117 ppm	140–460 ppm	0.04–0.22%	-	-	-
Ying et al. (2009)	10–350 ppm	300–1080 ppm	0.1–2.2%	-	-	-
Yap et al. (2005)	10–800 ppm	100–210 ppm	0.04–0.18%	-	-	-
Maurya and Agarwal (2014)	0–90 ppm	6.7–18.4 g/kWh	55–74 ppm	-	-	-
Mack et al. (2009)	0.5–8 ppm	100–1100 ppm	500–1400 ppm	-	10–19 ppm	-
Gharehghani (2018)	0.02–0.45 g/kWh	0.8–4.2 g/kWh	0.7–6.5 g/kWh	-	-	-

Shi et al. (2005) experimented with studying the effect of internal and external exhaust gas recirculation on the emission from the HCCI engine. Four-stroke single-cylinder, water-cooled, and naturally aspirated diesel engine was used. For internal EGR, at 1500 rpm, if IMEP is increased from 0.3 to 0.4 MPa, NO emission is less than 20 ppm at 10° CA valve overlap. If valve overlap changes to −20° CA valve overlap, NOx emission increased from 30 to 150 ppm. A drastic increase in NOx emission was observed at 40° CA negative valve overlap, and it reached 1100 ppm at 0.4 MPa IMEP. Maximum smoke opacity was observed at 0.4 MPa IMEP, and its value is 6.5% for 40° CA negative valve overlap, 6% for 20° CA negative valve overlaps, and 16% for 10° CA valve overlap. For External EGR, NOx emission in diesel engine was found to be approximately 360 ppm, but in the HCCI engine, NOx emission was reduced to 150 ppm at 0% EGR. As EGR rate is increased to 50%, NOx emission was further reduced to even less than 10 ppm. Maximum smoke opacity of 50% was observed for diesel engines at 40% EGR rate. In HCCI engine, smoke opacity was reduced to 10–30%. Maximum CO emission of diesel engine was 0.15% which was less than HCCI engines which produce CO emissions of 0.4%.

Singh and Agarwal (2016) examined HCCI combustion by using four test fuels. These test fuels are diesel, diesoline (15% v/v of gasoline with diesel), diesohol (15% v/v of ethanol with diesel), and diesosene (15% v/v of kerosene with diesel). A two-cylinder engine in which one cylinder operated on HCCI technique while the other on conventional compression ignition mode was used for experimentations. Compression ratio of HCCI cylinder was kept at 16.5. Analysis was also done for different air-fuel ratio.

CO emission was observed to be less than 10 g/kWh till "λ" value of 3.75 for all the fuels, but after 3.75, CO emission of diesohol increased rapidly to 50 g/kWh at "λ" 4.5 and 81 g/kWh at "λ" 5.25. CO emission for diesel, diesoline, and diesosene increased to 15 g/kWh, 20 g/kWh, and 19 g/kWh, respectively, at "λ" 4.5; and at "λ" 5.25, it further increased to 70 g/kWh, 65 g/kWh, and 80 g/kWh, respectively. HC emission also showed similar trend. Till "λ" 3.75, HC emission remained between 4 and 8 g/kWh and at "λ" 5.25, it increased to 12–20 g/kWh for all fuels. Maximum NOx emission found at "λ" 1.5 with a value of 2 g/kWh for diesosene, 1.5 g/kWh for diesel, 1.45 g/kWh for diesoline, and 1.2 g/kWh for diesohol. As "λ" value increased to 2.25, NOx emission reduced to 1.25 g/kWh for diesosene, 0.3 g/kWh for diesel, 0.45 g/kWh for diesoline, and 0.15 g/kWh for diesohol. If "λ" is further increased to 3.75, 4.5, or 5.25, NOx emission reduces to even less than 0.3 g/kWh. They figure out that addition of volatile additives like gasoline, kerosene, and alcohols to baseline diesel improved the emission characteristics in HCCI combustion, mainly up to medium range of loads. At higher range of loads, diesoline, and diesosene showed a higher knocking tendency when compared to diesel and diesohol.

Singh et al. (2014) experimented with studying the combustion, emission, and performance characteristics of the HCCI engine working on biodiesel using external mixture formation techniques. Experiments were performed on a four-stroke two-cylinder air-cooled direct injection diesel engine generator set which was run at a constant speed. One was operated in HCCI combustion mode out of two cylinders, while the other was operated on conventional CI combustion mode. Diesel and its various blends like B20 and B40 were used. Experiments were performed for different air-fuel ratio "λ" like 2.1, 2.7, 3.7, and 5.6. It was observed that CO and HC emissions were reduced by reducing the "λ" values or EGR rate. For diesel, maximum CO emission was 25 g/kWh for "λ" 5.6 and 30% EGR. For lower "λ" values of 3.7, 2.7, and 2.1, CO emission was less than 10 g/kWh. At "λ" 5.6 and 30% EGR, CO emission for B20 was 125 g/kWh; and for B40, it was 120 g/kWh. On reducing EGR to 0%, CO emissions reduced to 10 g/kWh for B20 and 85 g/kWh for B40. HC emission for diesel and B20 fuel ranges between 0 and 7.5 g/kWh for all values of λ and all EGR rates. But for B40 fuel, HC emissions were found to be 29 g/kWh for 30% EGR, 23 g/kWh for 15% EGR, and 25 g/kWh for 0% EGR at 5.6 "λ". For other values of "λ", HC emission ranges between 5 and 10 g/kWh. NOx emission was highest for diesel fuel at 2.1 "λ" around 33 g/kWh and 0% EGR, 23 g/kWh for 15% EGR, and 12 g/kWh for 30% EGR. As "λ" was increased to 5.6, NO

Table 2 Emission values for diesel

Emission parameters Authors	NO_X	HC	CO	CO_2
Singh and Agarwal (2012)	2–4 g/kWh	0.75–3 g/kWh	3–13 g/kWh	-
Gowthaman and Sathiyagnanam (2016)	10–600 ppm	50–100 ppm	0.05–0.37%	-
Shi et al. (2005)	10–1100 ppm	-	0.4%	-
Singh and Agarwal (2016)	0.15–2 g/kWh	4–20 g/kWh	10–80 g/kWh	-
Singh et al. (2014)	0–33 g/kWh	5–30 g/kWh	10–135 g/kWh	-
Patel and Yadav (2017)	10–120 ppm	22–120 ppm	0.01–0.09%	-
Mohanamurugan and Sendilvelan (2011)	20–500 ppm	20–140 ppm	0.1–0.3%	0.5–6.5%

Table 3 Emission values for petrol

Emission parameters Authors	NO_X	HC	CO
Hyvönen et al. (2003a, b)	-	70–110 g/kg	100–700 g/kg
Hyvönen et al. (2005)	0.3–0.65 g/kg	15–140 g/kg	20–40 g/kg
Hyvönen et al. (2003a, b)	0.2–0.37 g/kg	80–150 g/kg	80–550 g/kg

emission became almost negligible. Similar output was observed for B20 fuel also. For B40 fuel, NO emission was almost 0 for all vales of λ and EGR rates.

Patel and Yadav (2017) experimented with improving biodiesel and CNG-fuelled HCCI engine performance and emission characteristics. Experiments were carried out on a four-stroke single-cylinder high-speed water-cooled diesel engine which is coupled with an eddy current dynamometer. Thermocouples were fixed to measure temperatures. Gaseous emissions such as CO, HC, and NOx were measured using the exhaust gas analyzer connected to the exhaust pipe. It was seen that HC emission for all types of fuel increases with an increase in load. HC emission ranges between 42 and 76 ppm for diesel, 44 to 78 ppm for B5, 30 to 50 ppm for B10, 22 to 44 ppm for B15, and 80 to 120 ppm for CNG. On checking for NOx emission, it was least for CNG and ranges between 10 and 20 ppm. NOx emission for diesel and B10 first increases from 80 to 110 ppm if load was increased from 0 to 20% and then it decreases to 90 ppm at 40% and further increased to 120 ppm at 80%; and if load was increased to 100%, NOx emission decreased to 110 ppm. For B5, NOx emission increases from 30 ppm at 0% to 85 ppm at 100% and for B15, NOx emission increases from 70 ppm at 0% to 110 ppm at 100%. CO emission for diesel increases from 0.04% vol at 0% load to 0.09% vol at 100% load. Least CO emission was found for B15 fuel around 0.01–0.02% vol. For B5, B10, and CNG fuel, CO emission ranges between 0.02% vol and 0.08% vol. Hence, the best fuel for getting lease HC and CO emission is B15 and for getting the least NOx emission is CNG.

Mohanamurugan and Sendilvelan (2011) performed experiments to study the emission and combustion characteristics of different fuels in the HCCI engine. A four-stroke indirect injection diesel engine was used. Different modes of fuel injection were applied from the primary injector and pilot injector. HC emission decreases with an increase in load for HCCI mode while increasing load for conventional mode. Maximum HC emissions at 100% load for diesel and biodiesel were 110 ppm and 60 ppm, respectively. For diesel: diesel, biodiesel: biodiesel, diesel: biodiesel, biodiesel: diesel, diesel: petrol, and biodiesel: petrol, maximum HC emissions were 35, 25, 45, 140, 85, and 90 ppm. NOx emission for all modes increases with an increase in load. NOx emissions were around 500–550 ppm for conventional mode, much higher than for HCCI mode, around 300–400 ppm. CO emission was constant at 0.1% for all loads in the HCCI engine except in diesel: petrol mode. In diesel: petrol mode, CO emission was 0.3% for 0–20% load and 0.2% for 40–100% load. For diesel and biodiesel conventional mode, first increases to 0.5% and 0.4% till 40% load and then decreases to 0.3% and 0.2% at 100% load condition. Emission values for diesel are summarized in Table 2.

Petrol

Hyvönen et al. (2003a, b) experimented with studying the multi-cylinder HCCI engine operating range using the variable compression ratio. A 5-cylinder, 1.6-L engine with a variable compression ratio (VCR) range of 8:1 and 14:1 in the original was used. Its piston was modified to get VCR range of 9:1 and 21:1 and called it P21 piston. They compared the result with an earlier test of an HCCI engine having VCR range 9:1 and 17:1 and called it P17 piston. It can be seen that CO emission reduced by around 85% from 700 to 100 g/kg as the speed is increased from 1000 to 5000 rpm. Also, HC emission was reduced by around 35% from 110 to 70 g/kg as the speed increased from 1000 to 5000 rpm.

Furthermore, Hyvönen et al. (2005) performed a comparative study on operating conditions of HCCI engine using spark ignition and without using spark ignition. A 5-cylinder, 1.6-dm^3 engine having a variable compression ratio (VCR) ranging between 10:1 and 30:1 was used for the experiment. Approximately, standard valve timings were used, except the inlet valve closing, 10 crank angle degrees (CAD) earlier. NOx emission level was between 0.3 and 0.65 g/kg_{fuel} (4–8

Table 4 Emission values for LPG

Emission parameters Author	NO_X	HC	CO	CO_2
Kannan et al. (2020)	5–10 g/kWh	15–135 g/kWh	0.0443–0.4 g/kWh	5–11 g/kWh

Table 5 Emission values for different fuels

Emission parameter / Fuel	NO_X	HC	CO	CO_2	O_2	H_2O
Natural gas	0–800 ppm Yao et al. (2006) Olsson et al. (2003) Maurya and Agarwal (2014)	10–1100 ppm He et al. (2013) Polat (2015) Mack et al. (2009)	0.04–2.2% Yao et al. (2006) Ying et al. (2009) Yap et al. (2005)	4.76–6.3% Tsolakis et al. (2007)	12.6–10.8% Tsolakis et al. (2007)	5.02–6.4% Tsolakis et al. (2007)
	0–0.45 g/kWh Olsson et al. (2001) Gharehghani (2018)	0.8–18.4 g/kWh Olsson et al. (2001) Maurya and Agarwal (2014)	1–90 g/kWh Olsson et al. (2001) Gharehghani (2018)			
Diesel	10–1100 ppm Shi et al. (2005) Patel and Yadav (2017)	20–140 ppm Gowthaman and Sathiyagnanam (2016)	0.01–0.4% Shi et al. (2005) Patel and Yadav (2017)	-	-	-
	0.15–33 g/kWh Singh and Agarwal (2012) Singh and Agarwal (2016)	0.75–30 g/kWh Singh et al. (2014) Singh and Agarwal (2016)	3–135 g/kWh Singh and Agarwal (2012) Singh and Agarwal (2016)			
Petrol	0.2–0.65 g/kg Hyvönen et al. (2003a, b) Hyvönen et al. (2005)	70–150 g/kg Hyvönen et al. (2003a, b) Hyvönen et al. (2005)	20–700 g/kg Hyvönen et al. (2003a, b) Hyvönen et al. (2005)	-	-	-
Liquefied petroleum gas	5–10 g/kWh Kannan et al. (2020)	15–135 g/kWh Kannan et al. (2020)	0.0443–0.4 g/kWh Kannan et al. (2020)	5–11g/kWh Kannan et al. (2020)	-	-

ppm) in this test. NOx emission level is almost 0.1 g/kg_{fuel} (1–2 ppm) higher in spark-assisted HCCI combustion than normal HCCI emission at all air-fuel ratios. Engine out NOx emission increased as the air-fuel decreases. Unburned HC increased from 50 to 60 g/kg_{fuel} to a maximum of about 130–140 g/kg_{fuel} at around 60% initial slow heat release fraction and later decreased to stoichiometric 14–18 g/kg_{fuel} level. CO emissions are between 20 and 40 g/kg_{fuel} level.

Hyvönen et al. (2003a, 2003b) experimented on a supercharged HCCI engine to extend its operational range. A 5 cylinder 1.6-L engine with a variable compression ratio ranging between 9:1 and 21:1 was used for the experimentations. HC and CO emission decreased when throttle was provided. HC emission ranges between 100 and 150 g/kg at 1000 rpm and 80–100 g/kg at 2000 rpm. If catalyst is added, specific HC emission was reduced to less than 5 g/kWh. CO emission ranges between 100 and 450 g/kg at 1000 rpm and 80–550 g/kg at 2000 rpm. If the catalyst was added, specific CO emission was observed to be even less than 2 g/kWh. However, NOx emission increases on providing throttle. It increases from 0.2 to 0.37 g/kg at 2000 rpm. At 1000 rpm, NOx emission decreases from 0.25 to 0.2 g/kg. Emissions for Petrol fuel are compiled in Table 3.

Liquefied petroleum gas (LPG)

Kannan et al. (2020) performed a study to analyze LPG-fuelled diesel engines using HCCI combustion technique. Effect of preheating of intake air on emission of LPG-fuelled HCCI engine was analyzed under varying load conditions. Tests were conducted in an HCCI engine modified from single-cylinder diesel engine. Port fuel induction method was used to prepare a homogeneous charge. For the same load, HC and CO emissions were higher in the LPG-fuelled HCCI engine, whereas NOx and CO_2 emissions were lower than the conventional diesel engine. HC emission for diesel engine remained below 10 g/kWh, whereas for LPG, it reduced from 135 to 15 g/kWh on increasing the load. CO

emission for diesel engine was below 0.1 g/kWh, whereas for LPG, it reduced from 0.4 to 0.0443 g/kWh on increasing the load. NOx emission for diesel engine increased from 13 to 35.14 g/kWh for diesel engines, and for LPG, it ranged between 5 and 10 g/kWh. CO_2 emission of diesel engine was between 14 and 16 g/kWh, and for LPG-fuelled HCCI engine, it reduced from 11 to 5 g/kWh on increasing the load. Emissions from HCCI engines for LPG are summarized in Table 4.

Conclusion

Finally, emissions from HCCI engines for different fuels are summarized in Table 5.

Following are concluding remarks based on a critical review of emissions from HCCI using different fuels:

- NOx emission in HCCI engine is less than a conventional diesel engine, almost 1:10th the emission.
- Different fuels can be used in HCCI engines, and NOx emission is the least of HCCI engines worked on natural gas than diesel, petrol, or LPG.
- NOx emission is 90–95% lesser for natural gas-fuelled HCCI engine than diesel, petrol, or LPG, which will significantly help in reducing the environmental pollution.
- NOx emission increases by almost 0.17 g/kg_{fuel} as inlet pressure decreases by 0.4 bar. Also, NOx emission increases by 0.1 g/kg_{fuel} on an increment of speed by 1000 rpm.
- HC emission for natural gas-fuelled HCCI engine is 50–60% less than diesel or LPG, whereas CO emission is 90–95% less for LPG–fuelled HCCI engine.
- Exhaust gas recirculation also helps in reducing the NOx emission by around 300–400 ppm; however, HC and CO emission increases by 100–200 ppm.
- Advancing the combustion timing reduces the HC emission. If load increases in HCCI engine, HC emission gets reduced.

This review will be helpful for the researchers doing further studies on HCCI engine. Research can be done to solve various challenges involved in the implementation of HCCI engine in real life. Various performance parameters can also be analyzed.

Acknowledgements The authors are highly grateful to the Centre for Energy and Environment, Delhi Technological University, Delhi (India), for providing the basic facility to compile the work.

Author contribution Sanjeev Kumar Verma: methodology, writing—original draft. Subhashish Gaur: visualization. Tabish Akram: investigation. Samsher Gautam: supervision. Anil Kumar: conceptualization, Writing—review and editing, supervision.

Data availability Here, critical analysis has been done with reference to earlier research work. This is a kind of comprehensive review. Hence, there is no data used.

Declarations

Ethics approval and consent to participate Not applicable

Consent for publication Not applicable

Competing interests The authors declare no competing interests.

References

Gharehghani A (2018) Load limits of an HCCI engine fuelled with natural gas, ethanol, and methanol. Elsevier Ltd Fuel 239(2019):1001–1014. https://doi.org/10.1016/j.fuel.2018.11.066

Gowthaman S, Sathiyagnanam AP (2016) Effects of charge temperature and fuel injection pressure on HCCI engine. Elsevier Ltd. Alexendria Engineering Journal 55:119–125. https://doi.org/10.1016/j.aej.2015.12.025

Hasan MM, Rahman MM (2016) Homogeneous charge compression ignition combustion: advantage over compression ignition combustion, challenge and solution. Renew Sust Energ Rev 57:282–291

He BQ, Yuan J, Liu MB, Zhao H (2013) Combustion and emission characteristics of a n-butanol HCCI engine. Elsevier Ltd Fuel 115(2014):758–764. https://doi.org/10.1016/j.fuel.2013.07.089

Hussaini SY, Lahane S, Patil NG (2016) Analysis of performance and emission characteristics of homogeneous charge compression ignition (HCCI) Engine. Procedia Technology 25(2016):854–861

Hyvönen J, Haraldsson G, Johansson B (2003a) Operating range in a multi cylinder HCCI engine using variable compression ratio. Society of Automotive Engineer JSAE 20030178, SAE 2003-01-1829

Hyvönen J, Haraldsson G, Johansson B (2003b) Supercharging HCCI to extend the operating range in a multi-cylinder VCR-HCCI engine. SAE Technical Paper Series 2003-01-3214

Hyvönen J, Haraldsson G, Johansson B (2005) Operating conditions using spark assisted HCCI combustion during combustion mode transfer to SI in a multi-cylinder VCR-HCCI Engine. SAE Technical Paper Series 2005-01-0109

Janathanan K, Pandian V, Dhandapani M, Kanagara K (2016) A performance, combustion and emission study on HCCI engine: trends and innovations. SAE Technical Paper 2016-28-0013

Kannan S, Mahalingam S, Srinath S, Sivasankaran M, Kannan S (2020) An experimental study in HCCI combustion of LPG in diesel engine. 37:3625–3629. https://doi.org/10.1016/j.matpr.2020.09.774

Mack JH, Aceve SM, Dibble RW (2009) Demonstrating direct use of wet ethanol in a homogeneous charge compression ignition (HCCI) engine. Energy 34(2009):782–787. https://doi.org/10.1016/j.energy.2009.02.010

Maurya RK, Agarwal AK (2010) Experimental study of combustion and emission characteristics of ethanol fuelled port injected homogeneous charge compression ignition (HCCI) combustion engine. Elsevier Ltd. Appl Energy. https://doi.org/10.1016/j.apenergy.2010.09.015

Maurya RK, Agarwal AK (2014) Experimental investigations of performance, combustion and emission characteristics of ethanol and methanol fuelled HCCI engine. Elsevier Ltd Fuel Processing

Technology 126(2014):30–48. https://doi.org/10.1016/j.fuproc.2014.03.031

Mohanamurugan S, Sendilvelan S (2011) Emission and Combustion Characteristic of Different Fuels in a HCCI Engine. International Journal of Automotive and Mechanical Engineering ISSN 1985-9325 (Print) ISSN 2180-1606 (Online) 3:279–292

Mulane SA, Limaye SD (2016) Homogeneous charge compression ignition (HCCI) engine technology – a review. INPRESSCO E-ISSN 2277–4106, P-ISSN 2347–5161

Ohur EO, Kariuki SM (2014) Effect of car emission on human health and the environment. Int J Appl Eng Res 9(21):11121–11128

Olsson JO, Tunestal P, Haraldsson G and Johansson B (2001) A turbo charged dual fuel HCCI engine. SAE Technical Paper Series 2001-01-1896

Olsson JO, Tunestal P, Ulfvik J, Johansson B (2003) The effect of cooled EGR on emissions and performance of a turbocharged HCCI engine. SAE Technical Paper Series 2003-01-0743

Patel DS, Yadav M (2017) Experimental study to improve the performance and emission characteristics of HCCI engine using blends of bio-diesel and CNG. IRJET e-ISSN: 2395-0056 p-ISSN 2395-0072

Polat S (2015) An experimental study on combustion, engine performance and exhaust emissions in a HCCI engine-fuelled with diethyl ether–ethanol fuel blends. Elsevier Ltd Fuel Processing Technology 143(2016):140–150. https://doi.org/10.1016/j.fuproc.2015.11.021

Shi L, Cui Y, Deng K, Peng H, Chen Y (2005) Study of low emission homogeneous charge compression ignition (HCCI) engine using combined internal and external exhaust gas recirculation (EGR). Elsevier Ltd Energy 31(2006):2665–2676

Singh AP, Agarwal AK (2012) An experimental investigation of combustion, emissions and performance of a diesel-fuelled HCCI engine. Society of Automotive Engineer Paper No. 2012-28-0005

Singh AP, Agarwal AK (2016) Diesoline, diesohol and diesosene-fuelled HCCI engine development. Journal of Energy Resources Technology. https://www.researchgate.net/publication/302870619

Singh G, Singh AP, Agarwal AK (2014) Experimental investigations of combustion, performance and emission characterization of biodiesel-fuelled HCCI engine using external mixture formation technique. Sustainable Energy Technologies and Assessments 6(2014):116–128. https://doi.org/10.1016/j.seta.2014.01.002

Tsolakis A, Megaritis A, Yap D (2007) Application of exhaust gas fuel reforming in diesel and homogeneous charge compression ignition (HCCI) engines fuelled with bio-fuels. Elsevier Ltd. Energy 33(2008):462–470

Venkataramana P (2013) Homogeneous charge compression ignition (HCCI) engine. International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181, Vol. 2 Issue 2

Yao M, Chen Z, Zheng Z, Zhang B, Xing Y (2006) Study on the controlling strategies of homogeneous charge compression ignition combustion with fuel of dimethyl ether and methanol. Elsevier Ltd 85(2006):2046–2056

Yap D, Peucheret SM, Megaritis A, Wyszynski ML, Xu H (2005) Natural gas HCCI engine operation with exhaust gas fuel reforming. Elsevier Ltd International Journal of Hydrogen Energy 31(2006): 587–595

Ying W, Li H, Jie Z, Longbao Z (2009) Study of HCCI-DI combustion and emissions in a DME engine. Elsevier Ltd Fuel 88(2009):2255–2261. https://doi.org/10.1016/j.fuel.2009.05.008

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

PAPER

Evaluation of the residual stress of HVOF sprayed carbon coating after wear testing conditions using ANN coupled Taguchi approach

To cite this article: Ankit Tyagi *et al* 2021 *Surf. Topogr.: Metrol. Prop.* **9** 035027

View the article online for updates and enhancements.



IOP ebooks™

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection–download the first chapter of every title for free.

This content was downloaded from IP address 14.139.251.100 on 31/08/2021 at 10:28

Surface Topography: Metrology and Properties

PAPER

Evaluation of the residual stress of HVOF sprayed carbon coating after wear testing conditions using ANN coupled Taguchi approach

Ankit Tyagi[1,2,*], **Qasim Murtaza**[2] and **R S Walia**[3]

[1] Department of Mechanical Engineering, SGT University, Haryana, India
[2] Department of Mechanical Engineering, Delhi Technological University, Delhi, India
[3] Department of Production & Industrial Engineering, PEC, Chandigarh, India
[*] Author to whom any correspondence should be addressed.

E-mail: tyagiankit10506@yahoo.com

RECEIVED
12 April 2021

REVISED
8 August 2021

ACCEPTED FOR PUBLICATION
19 August 2021

PUBLISHED
27 August 2021

Keywords: numerical modeling, artificial neural network (ANN) model, Taguchi L9, analysis of variance (ANOVA), HVOF coating

Abstract

The as deposited carbon coating was successfully synthesized from agricultural waste using High Velocity Oxy-fuel (HVOF) thermal spray deposition technique. The surface topography of as deposited carbon coating, confirms the presence of molten, semi-molten, and un-melted grains along with the formation of lamellae. Taguchi L_9 orthogonal Array (OA) and Artificial Neural network (ANN) models were used to predict and optimize the process parameters for residual stress after wear testing conditions. The estimated mean (μ) obtained from optimal responses is 11 MPa. Three confirmation tests of as-deposited carbon coating for residual stress after wear testing conditions were conducted for optimal level of parameters at level 3 of Temperature (250 °C), level 3 of sliding velocity (0.3 m s^{-1}), and level 3 of load (80 N). The percentage (%age) contribution of as-deposited carbon coating for residual stress after wear testing conditions from Analysis of Variance (ANOVA) prediction was highest for temperature (51.23%), followed by 32.8% of sliding velocity, and 12.5% for a load.

1. Introduction

In the Modern Era, additive manufacturing, welding, and coating are some of the widely used key manufacturing processes. The metallic component processing often trailed to inhomogeneous input of thermal heat, which eventually results residual stresses formation. Residual stress's magnitude and distribution plays a vital role in designing and manufacturing of correct product as residual stresses either have a negative or positive effect on the wear resistance, corrosion, and load bearing capabilities [1]. The residual stresses formation may cause macroscopic cracking or even delamination of thermal spray coating [2]. Welding is primarily categorized by thermal distortion and residual stress may hurt welding due to alternating load resistance. Similarly, the additive manufacturing process may have both combined effects due to thermal distortion and unexpected failure of components due to multiaxial stress formation [3, 4]. Residual stresses, corrosion, and fatigue are prominent challenges in the aerospace, automotive, marine, and chemical industries [5–7]. Raghupathy *et al* [8] and Ma *et al* [9] observed that HVOF, cold spray, and plasma arc likes coating techniques have been employed for wear and corrosion resistance. Moreover, HVOF thermal spray technique has been widely used to enhance performance properties like wear and corrosion resistance [10] and also used to produces a good interface with high quality and low porosity coating. The experimental studies [7, 11, 12] reported that the porosity of Inconel 625 reduces less than 1%, while coating hardness increase from 200 to 450 HV. HVOF thermal spray techniques provide low-cost optimum adhesion for wear and corrosion resistance of carbide (WC-Co) coatings [13]. The experimental study of Padilla *et al* [14], reported that compressive residual stresses, coating bond, and fatigue strength correlate all. The experimental study reported that wear behavior plays an important role in residual stresses. The study shows that the percentage change in residual stress is between 8.64 to 13.54 MPa after wear test [15]. The study also concluded that temperature had highest percentage contribution from Analysis of Variance (ANOVA) prediction. Another experimental study shows that compressive residual stresses are developed ranging

© 2021 IOP Publishing Ltd

from −109 to −5 MPa on the coating samples after wear test [15]. The experimental study of Tyagi *et al* [16] shows that the residual stresses decreases as the temperature, load and sliding velocity increases during wear testing conditions.

Temperature, load and sliding velocity are the three major factors that influencing residual stresses, Coefficient of COF, and wear of developed coating, and hence optimizing these parameters through ANN, Fuzzy logic, Root Square method (RSM), Genetic algorithm, particle swarm optimization, etc, will certainly help in achieving admissible results [17]. These optimizing techniques are exhaustively used to optimize vital parameters in material science, to reduce any sort of wastage and improve the quality of materials used for engineering applications. Initially, the Taguchi design method is used to identify the optimum process parameters and eminent the critical factors to estimate the influence of process parameters on HVOF coating [18]. However, due to the combined effect of many process parameters of HVOF coating, it may fail to give precise results. In recent years, many researchers are focused upon ANN optimization, to solve complex problems. It is an effective tool for prophesying the output generation for process parameters of demonstrative input sets [19]. ANN is one of the powerful modeling tools [20], used for modeling with different output and input parameters using optimum network algorithms.

In the Modern Era, processing of metallic components often trailed to inhomogeneous input of thermal heat, which eventually results in the formation of residual stresses. Magnitude and distribution of residual stresses play a vital role in designing and manufacturing of correct product. It is quite evident from the literature that there is a need to develop coating which can resist residual stress formation within the component. The study shows microstructural analysis, impact of particle velocity and optimizing for residual stress after wear testing conditions using Taguchi L_9 OA and the ANN model.

The present research address the following questions:

1. Deposition of carbon coating using HVOF process.
2. To study surface morphology and elemental analysis of as-deposited coating.
3. To study the impact of particle velocity on residual stresses using finite element analysis (ABAQUS software).
4. Optimizing the experimental results for residual stress after wear testing conditions using Taguchi L_9 OA and the ANN model.

2. Experimental procedure

2.1. Coating deposition

Good eminence carbon coatings from agricultural waste were synthesized and developed on the steel plate, using the HVOF process. Agricultural waste was initially pyrolyzed into the thermal furnace at 1 MPa and 550 °C temperature. The produced activated carbon from waste and other carbon powders were mixed and blend in the ball mill to get carbon powder with a reduced size of 50 ± 5 μm. The physicochemical properties of the produced powder are the grain size, density, and melting point powder was 50 μm, 2.10 g cm^{-3}, and 3500 °C ± 100 °C respectively [21]. Sieve Grain Size Analysis was done in order to determine the grain while density of particle was measured using bulk density tester or density analyzer. The melting point of powder was determine by using determine using calorimeter. The produced activated carbon from agricultural waste is having high commercial values and are having high carbon contents, high density, and good mechanical properties.

Initially, 80 × 80 × 5 mm size steel plate was ground using 180–2000 grit emery paper, to get the smooth surface. After that, the PR-9182 model pressure blaster, made by Metallizing Equipment Co. Pvt. Ltd, Jodhpur, INDIA, was used to prepare and deposit SiC additive layer, while the dust generated in the blast cabinet as a result of the blasting operations is collected in a dust collector. The SiC additive layer was used for intensification of adhesive strength among substrate and coating material. The HIPOJET-2700M MEC-made HVOF system was used for spraying carbon powder at supersonic speed jet in HVOF thermal spray process. During the growth of carbon coating, N_2 gases was continuously supplied on steel plate and powder spray system, in order to prevent coating from contamination. The coating was deposited using 4 kg h^{-1} powder feed rate, treatment time of 10 min and deposition efficiency of 55%. The particle velocity was about 650 m s^{-1} and was measured using spray watch system associated with high power Laser pulsed and CCD camera sensor. Table 1 demonstrate the deposition conditions used to grow HVOF coating from agricultural waste. Figure 1 shows the flow chart of experimental procedure followed in the present research of as deposited coating. It is extended the matrix method of structural analysis to 2D continuum domains by discretizing the domain into triangular elements and stiffness matrices of an element were obtained from the strain energy expressions by assuming the linear variation of displacement over the element. Figure 2 shows the stages of finite Element Analysis.

2.2. Coating characterization and residual stress testing

The surface morphology, elemental and microstructural composition of developed coatings was characterized by FESEM attached with EDS available at IUAC, Delhi, India. The crystal structure and phase identification of coatings were characterized and

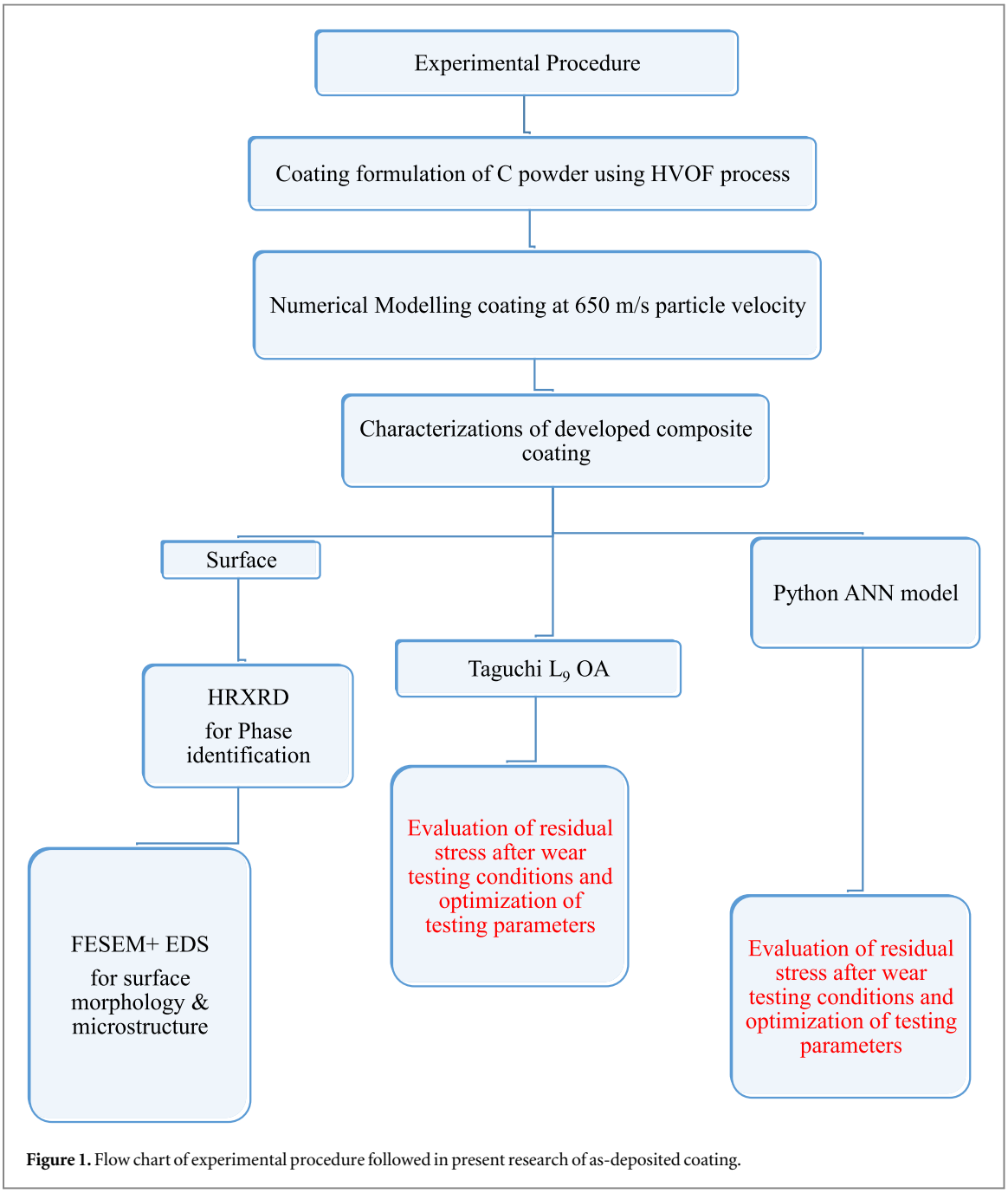


Figure 1. Flow chart of experimental procedure followed in present research of as-deposited coating.

Table 1. Details of coating deposition parameters.

No.	Parameters used	Standards
1.	Coating Composition	agricultural waste produced carbon powder
2.	Deposition technique	HVOF
3.	Substrate material	Mild Steel plate
4.	Surface pre-treatment	Acetone
5.	Powder feed rate	4 Kg h^{-1}
6.	Carrier gases	N_2
7.	Treatment time	10 min
8.	Deposit efficiency (%)	55
9.	Fuel flow rate	10 ± 1 (/h

confirm by HRXRD (PANalytical systems) using Cu anode material, generator voltage 45 KV and tube current of 40 mA in conventional Bragg mode.

The μ-X360 residual stress analyzer (Pulstec manufacturer) was used to evaluate the residual stresses developed on the carbon coatings. The residual stress was examined using the cos(α) method based on Braggs law with x-ray indentation of 45° and support of sensor unit. The residual stress is measured after wear testing at different conditions. The pin-on-disk tribometer (TR-20L-PHM800-DHM850) was used wear testing at different conditions of carbon coatings. The maximum temperature range of the rotary tribometer was from ambient to 850 °C, applied load from 20 to 200 N, and sliding velocity from 0.1 to 15 m s^{-1}. The experimental tests were conducted to a 1500 m travel distance without lubrication, G99 grade pin.

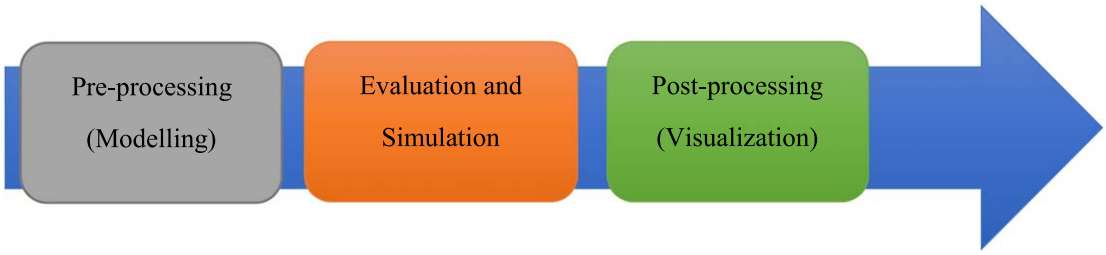


Figure 2. Stages of Finite Element Analysis.

2.3. Numerical approach using finite element analysis

To investigate the coating behavior, there are several analytical approaches however, under plastic deformation ABAQUS gives better results. It is a general Finite Element Manufacturing (FEM) goal that utilizes an implicit integration scheme, explicit integration schemes to solve extremely non-linear device communication under transient loads, and offers advanced computational fluid dynamics (CFD) capabilities and advanced electromagnetic computational problems. It is extended the matrix method of structural analysis to 2D continuum domains by discretizing the domain into triangular elements and stiffness matrices of an element were obtained from the strain energy expressions by assuming the linear variation of displacement over the element.

There are different modules in this numerical approach (1) property, (2) part module, (3) interaction, (4) step module, (5) assembly, (6) boundary conditions, (7) mesh generation, and (8) optimization. The material module followed the Johnson cook equation:

$$\sigma = (A + B\epsilon_p^n)(1 + C \ln \epsilon^*)(1 - (T^*)^m) \quad (1)$$

$$T^* = \frac{T - T_0}{T_m - T_0} \quad (2)$$

Where A is yield stress, B is work-hardening parameter, and C is dimensionless strain rate parameter, respectively, and T^* is the dimensionless unit. Also, T_m and T_0 are the melting temperatures and reference temperature respectively. Different boundary conditions were used in the ABAQUS/Explicit for the numerical analysis.

Many researchers numerically investigated the variation of coating and substrate materials to analyses variation in the thermal and mechanical properties [22–24]. The Mild steel substrate has coated with carbon powder sprayed through HVOF coating that has been also numerically investigated in this study. In the current research, the carbon particles are deposited at a velocity of 650 m s^{-1} on the steel substrate. The physicochemical properties of the produced powder are the grain size, density, and melting point powder was 50 μm, 2.10 g cm^{-3}, and 3500 °C ± 100 °C respectively [21]. The stress-variation, deformation pattern contour formation for the multi-particle impact, were shown in the figure 3(a) Compressive residual stress, and (b) Von-Mises stress value (c) residual stress analysis (d) Coating analysis at the velocity of 650 m s^{-1}. The numerical modeling results confirm that HVOF thermal spray coating promotes compressive residual stresses in the as-deposited coating due to high particle velocity and differences in thermal temperatures between coating and substrate materials. Particle velocity plays an important role in producing mechanical locking bonds between splats and substrate material. Good bonding of thermal spray coating after deposition leads to compressive residual stresses [25].

2.4. Design of ANN model to predict residual stress

In the present work, Python and Keras were employed and the ANN model with Adams optimizer and Google Colab for computation was accomplished as shown in figure 4. The model was trained with 27 experimental values, while input and output values are normalized between −1 to 1, given all input parameters as equal importance to the response. ANN comprises mainly 3 layers, an input layer, an output layer, and a hidden layer. The input layer collects the required input data and transferred it to the hidden layer, which finally covers the input signals to output through nodes. In the hidden layer, every node has links, which are connected to both the input and output layer [17].

To accomplish an error function to minimize error or loss equation was used as follows.

$$L(\mathrm{Y}, \acute{\mathrm{Y}}) = -\frac{1}{N}\sum_{n\epsilon N}\sum_{i\epsilon C} Y(n, i)\log(\acute{Y}n, \mathrm{i}) \quad (3)$$

The prediction of ANN model can be done using forward propagation. We can calculate the prediction using following equations:

$$z_1 = xW_1 + b_1 \quad (4)$$

$$a_1 = \tanh(z_1) \quad (5)$$

$$z_2 = a_1W = +b_2 \quad (6)$$

$$a_2 = \mathrm{softmax}(z_2) \quad (7)$$

The prediction of the ANN model can be done using backward propagation. We can calculate the prediction using the following equations:

$$\delta_3 = \acute{\mathrm{Y}} - \mathrm{Y} \quad (8)$$

$$\delta_4 = (1 - \tanh^2 z_1) o \delta_3 W_2{}^{\mathrm{T}} \quad (9)$$

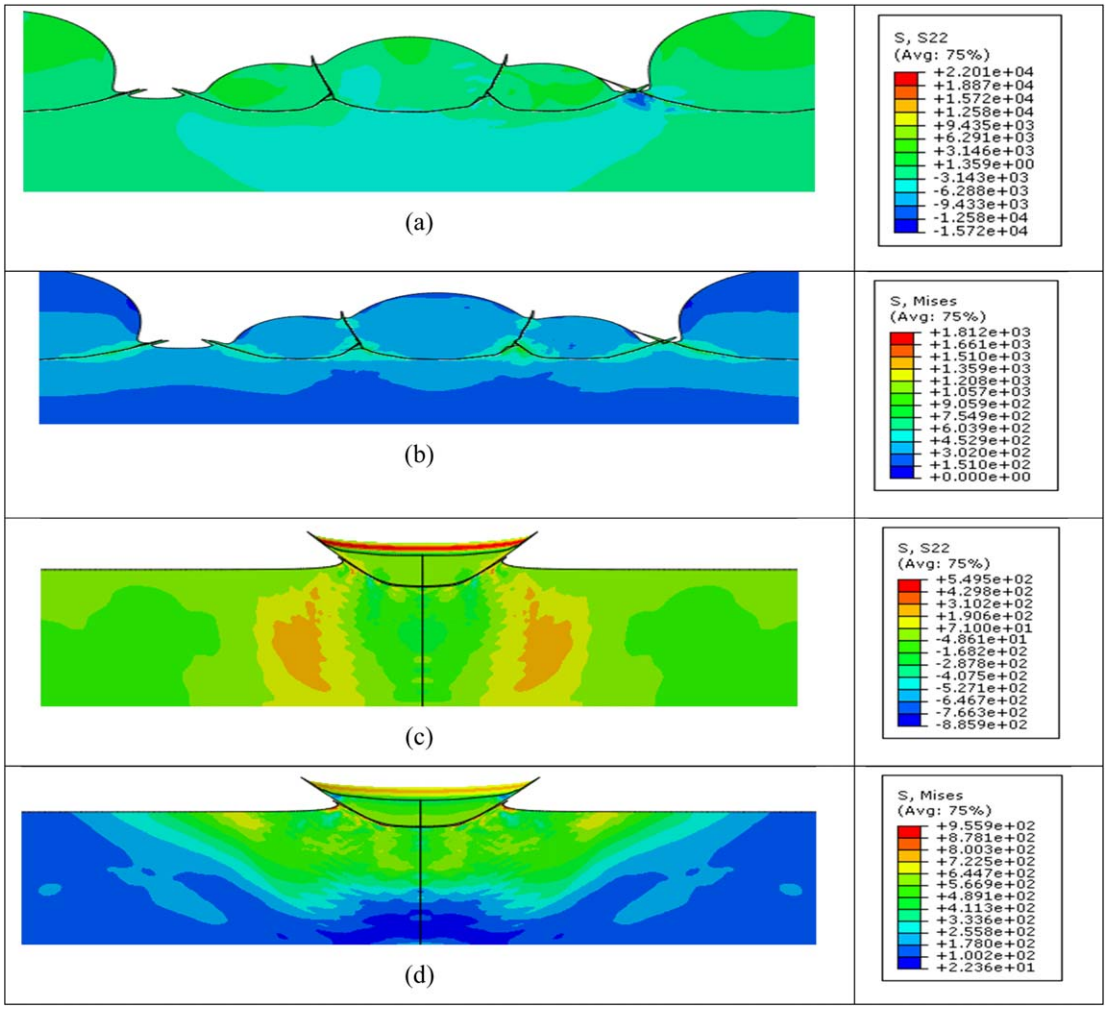


Figure 3. (a) Compressive residual stress and (b) Von-Mises stress value (c) residual stress analysis (d) Coating analysis at the velocity of 650 m s^{-1}.

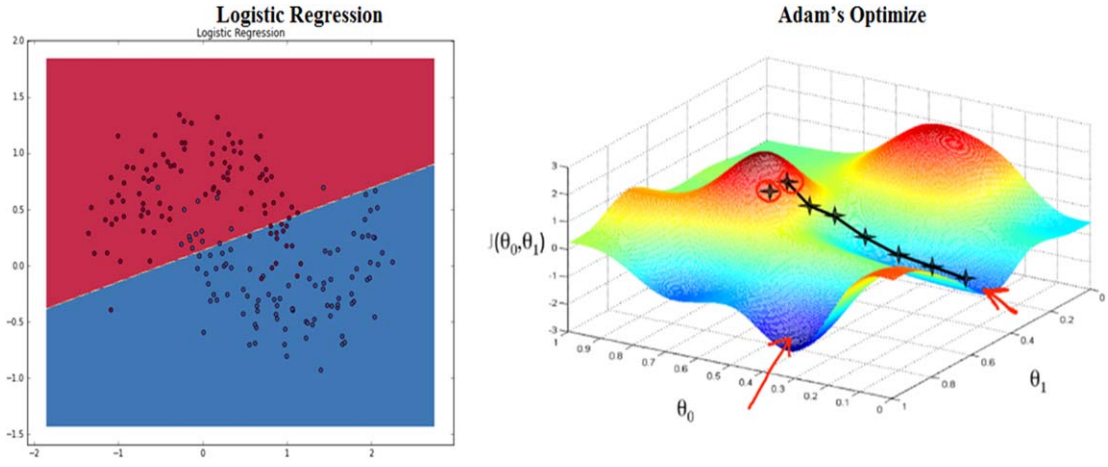


Figure 4. Residual stress values for ANN model using Adams optimizer.

$$\frac{\partial L}{\partial W2} = \alpha_1{}^{\mathrm{T}}\delta_3 \tag{10}$$

$$\frac{\partial L}{\partial b2} = \delta_3 \tag{11}$$

$$\frac{\partial L}{\partial W1} = \mathrm{x}^{\mathrm{T}}\delta_2 \tag{12}$$

$$\frac{\partial L}{\partial b1} = \delta_2 \tag{13}$$

Table 2. DOE of as-deposited carbon coating chosen for the Taguchi's method.

Parameters	Factors	Level 1	Level 2	Level 3
Temp. (C)	A	50	150	250
Sliding Velocity (m s^{-1})	B	0.1	0.2	0.3
Load (N)	C	60	70	80

2.5. DOE using Taguchi L_9 OA

In the present work, Design of Experiment using Taguchi L_9 OA was employed to predict the optimum value of as-deposited carbon coating for residual stress. Table 2 shows the design of experiments of as-deposited carbon coating for residual stress chosen for Taguchi's method. Presently, three process parameters were analyzed at three different levels with a total of 9 different compositions. All the compositions were tested thrice, so a total of 27 Taguchi L_9 OA test results of deposited carbon coating were obtained to analyze the effects of three factors (parameters) at three different levels on residual stress as shown in tables 2 and 3.

Lower the better form prominence physical characteristics were used for scrutiny S/N ratio of as-deposited carbon coating for residual stress. In the Taguchi method, the S/N ratio was used to recognize the controlling factors that reduce inconsistency in output. The S/N ratio prominence characteristics form may be presumed in the equation [26, 27]. ANOVA was instigated to identify statistically significant parameters of as-deposited carbon coating for residual stress experimental test results and further optimum parameters were predicated through S/N ratio and ANOVA analyses.

$$(S/N)_{LB} = -10 \log\left\{\frac{1}{T}\sum_{T}^{j=1}(X2j)\right\} \quad (14)$$

2.6. ANN model

The present developed ANN model constitutes 3 layers:

1. Input layer: Temperature, sliding velocity, and load as input nodes,
2. Hidden layer with R12 and R10 nodes,
3. Output layer with residual stress as an output node.

The literature [28–30] also confirms that at least 7 hidden nodes should be there for the ANN model. Hence in present research an ANN model using Python with feedforward and backward propagation and network topology 3-12-10-1 demonstrated as figure 5 is worthy, thereby topology 3-12-10-1 stands for an ANN model that was developed with 3 input nodes, 12 and 10 hidden nodes and a single output node.

ANN comprises mainly 3 layers, an input layer, an output layer, and a hidden layer. The input layer collects the required input data and transferred it to the hidden layer, which finally covers the input signals to output through nodes. The architecture of the ANN model contained only 2 layers with 64 neurons each and one flattening layer. Test/Train split was used for validation. 20% of the total data used 0.013d for testing and validation. The rest of the 80% was used for training. Optimizer ran for 100 epochs in a single batch. The saturation and fluctuation of loss in later epochs can be attributed to fluctuating Residual Stress values for the same input values (Temperature, SV, and load) as a function of time.

3. Results and discussion

3.1. Coating characterization

Figures 6(a), (b) shows the FESEM results of carbon coating. The deposited carbon show presence of molten, semi-molten, and un-melted structured grains [21]. Figures 6(c), (d) shows the surface morphology and EDS result of carbon coating having weight % of elements present in the coating are 80% carbon, 15% oxygen, and rest weight % of aluminum (1.65%), iron (0.65%), and tungsten (1.60%). Figure 6(e) shows the HRXRD spectra and shows carbon peaks at $2\Theta = 13.9°$, $2\Theta = 17.2°$, $2\Theta = 26.5°$, $2\Theta = 44.6°$ and at $2\Theta = 44.6°$ respectively were observed on developed coating. The XRD peak at $2\Theta = 26.5°$ of lattice plane (002) can be assigned to hexagonal graphite (PDF, 01-071-3739), at $2\Theta = 44.6°$ and $82.6°$ witnessed the presence of hexagonal carbon phased (DLC) (PCPD-79-1471). The other peak around $2\Theta = 17.2°$, witnessed hexagonal SiC of lattice plane (002) (PCPD-73-1749).

3.2. Effect of processing parameters using ANN model on residual stress

Figure 7 shows the 3D Debye ring, distortion ring, and profile ring of tested samples. The results of ANN prediction as in figure 8, show the deviation of residual stress concerning temperature on the as-deposited carbon coating. The experimental results show that temperature exerts more contribution as also confirmed from ANOVA in controlling the residual stresses of the deposited coating. It was observed that with increase in temperature, the residual stresses on the sample decrease due to decrease in thermal gradients while up to 400 °C temperature, structural stresses remain stable [31–33].

From figure 8 it is quite clear that residual stresses decrease as the velocity fluctuates from 0.1–0.3 m s^{-1} due to change in grain structure and state of stress-free condition may be established as the test conditions (sliding velocity) increases [14, 34, 35]. The lowest residual stress value was attained at 0.3 m s^{-1} velocities, while the highest at 0.1 m s^{-1} velocities. The sliding velocity emerged as the second most contributing factor as also confirmed from ANOVA in controlling the residual stresses of the deposited coating.

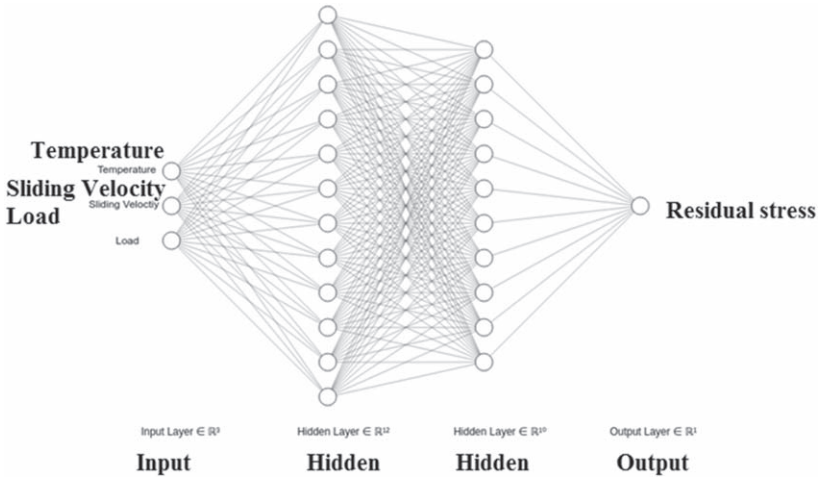


Figure 5. Residual stress values for ANN model developed.

Table 3. Taguchi L_9 OA test results of as-deposited carbon coating for residual stress.

Composite compositions	Temperature (C)	Sliding velocity (m s^{-1})	Load (N)	Residual Stress (compressive) Test 1(T1)	Test 2 (T2)	Test 3 (T3)	S/N ratio (DB)
1	50	0.1	60	210	205	200	−46.44
2	50	0.2	70	140	150	145	−43.33
3	50	0.3	80	90	85	95	−39.08
4	150	0.1	70	140	145	135	−42.92
5	150	0.2	80	110	105	115	−40.82
6	150	0.3	60	100	90	95	−39.41
7	250	0.1	80	70	75	65	−36.90
8	250	0.2	60	90	80	100	−39.08
9	250	0.3	70	40	35	45	−32.04

T = overall mean of residual stress (Compressive) = 110 MPa.

In this investigation, experimental results show that residual stresses obtained at 80 N were lowest on comparing with 60 N and 70 N load for as-deposited carbon coating due to change in grain structure and state of stress-free condition may be established as the test conditions increases [14, 33–35]. The results of ANN prediction as in figure 8 show the deviation of residual stress concerning load on the as-deposited carbon coating. The load exerts the third most important contributing factor as also confirmed from ANOVA in controlling the residual stresses of the deposited coating.

3.3. Effect of processing parameters using Taguchi model

3.3.1. Main effect of of as-deposited carbon coating

Table 4 shows the main effects of as-deposited carbon coating for residual stress and calculated all the responses of S/N ratio at a different level and plotted in figure 9.

Figure 9(a) shows the deviation of residual stress concerning temperature on the as-deposited carbon coating. The experimental results show that residual stresses attained were lowest at 250 °C on relating with 50 °C and 150 °C temperature for as-deposited carbon coating. The S/N ratio of as-deposited carbon coating at 250 °C temperature was highest on equating with 50 °C and 150 °C temperature. The above experimental result shows that residual stresses were lower at 250 °C temperature for as-deposited carbon coating and the larger value of S/N ratio proposes less noise and strong signal, therefore offers optimal results. The residual stress decreases as the operating temperature increases due to a reduction in thermal stresses while up to 400 °C temperature, structural stresses remain stable [31] or due to lower thermal expansion [14, 34].

From figure 9(b) it is quite clear that residual stresses decrease as the velocity varies from 0.1–0.3 m s^{-1}. At 0.3 m s^{-1} velocities, lowest residual stress value was obtained, while at 0.1 m s^{-1} velocities, highest residual stress was attained. However, the S/N ratio of as-deposited carbon coating at 0.3 m s^{-1} was highest on comparing with other test condition (Sliding velocity). From the above experimental results, it was observed that the S/N ratio increases as the sliding velocity decreases and was the highest corresponding to

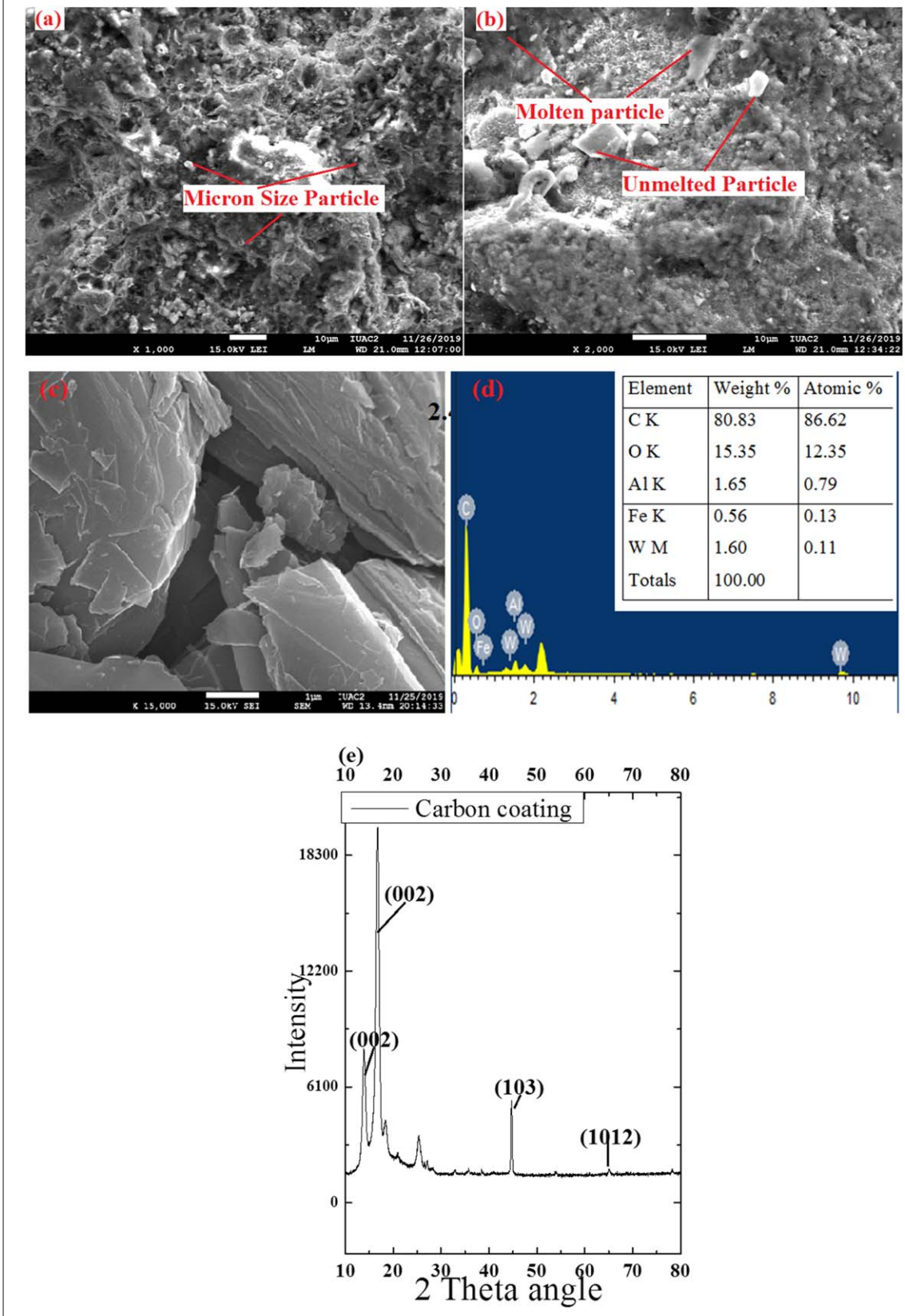


Figure 6. (a)–(d) FESEM + EDS image (e) HRXRD of carbon coating.

0.3 m s^{-1} sliding velocity. Similar experimental results were quoted by Bouzakis *et al* [33] and stated that change in grain structure and state of stress-free condition may be established as the test conditions (sliding velocity) increases [14, 34].

Figure 9(c) shows the deviation of residual stress concerning load on the as-deposited carbon coating. The experimental results show that residual stresses obtained at 80 N were lowest on comparing with 60 N and 70 N load for as-deposited carbon coating.

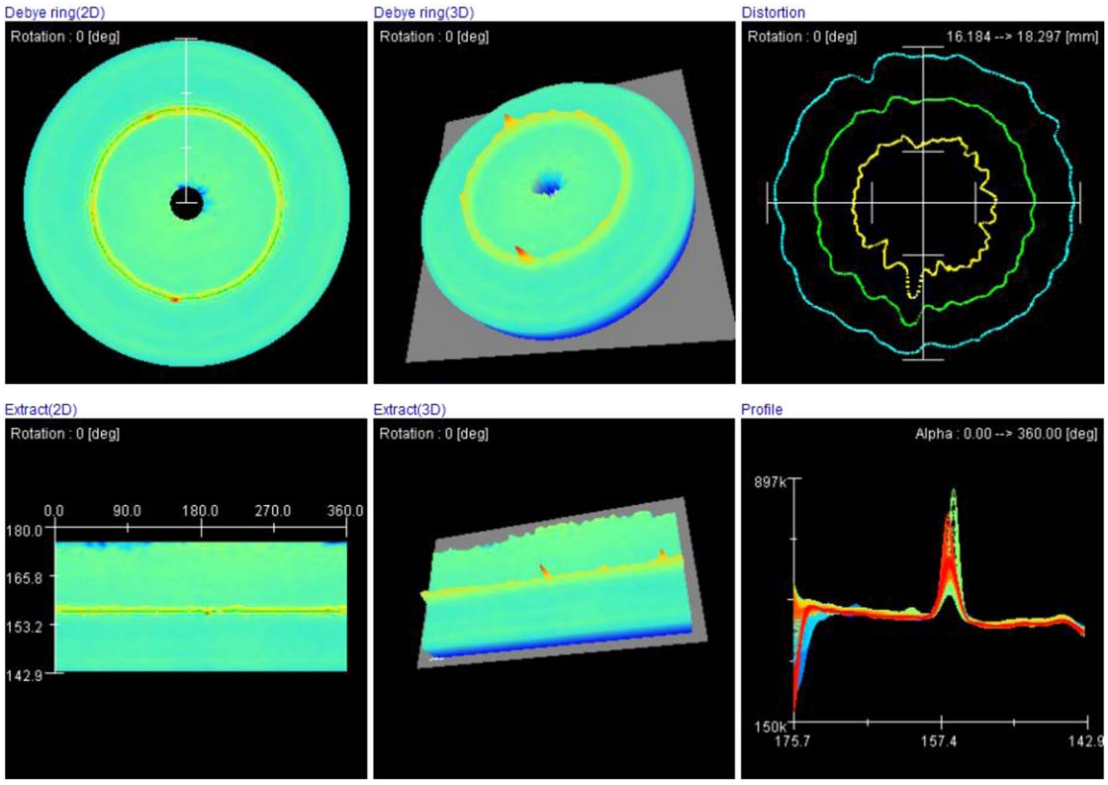


Figure 7. 3D Debye ring, distortion ring, and profile ring after the residual stress test.

However, the S/N ratio of as-deposited carbon coating was highest at 80 N load on comparing with 60 N and 70 N load. The above experimental result shows that residual stresses were lower at 80 N load for as-deposited carbon coating and the superior value of S/N ratio suggests less noise and strong signal, therefore offers optimal results. Similar experimental results were quoted by Bouzakis *et al* [32] and stated that change in grain structure and state of stress-free condition may be established as the test conditions increases [14–16,34].

In the current research, lower the better type of characteristic was used, for better results for as-deposited carbon coating was chosen, lower values of residual stress were sought. Figure 9 shows that optimum results of deposited carbon coating for residual stress were obtained at level 3 of parameters A, B, and C. It was further noted that the highest value of the S/N ratio corresponds to the lowest value of mean response for parameters A, B, and C. The experimental result of Bouzakis *et al* [33] stated that a change in grain structure and state of stress-free condition may be established as the test conditions increases. The residual stress decreases as the operating temperature increases due to reduction in thermal stresses while up to 400 °C temperature, structural stresses remain stable on as-deposited carbon coating [14–16,34], similar results were also obtained in the present research. The distortion images obtained after residual stress testing show no structural changes with less or no damage to the surface as shown in figure 6.

3.3.2. Pooled ANOVA of as-deposited carbon coating

To examine the significance of residual stress testing parameters, ANOVA was performed. Table 5 shows the pooled ANOVA raw and S/N data of residual stress of as-deposited carbon coating. It was evident that mean & variation in residual stresses are considerably affected by the process parameters. The % age contribution of as-deposited carbon coating for residual stress was highest for temperature (51.23%), followed by 32.8% of sliding velocity, and 12.5% for a load.

3.3.3. Estimation of optimal characteristics

3.3.3.1. Optimal response characteristics for residual stress (predicted mean)

Table 6 shows average residual stress values of as-deposited carbon coating at the optimum level of significant factors. The estimated mean of as-deposited carbon coating for residual stress at an optimal level can be obtained from equation (15) [28–33,35].

$$\mu_{RS} = A_3 + B_3 + C_3 - 2T \tag{15}$$

$$\mu_{RS} = 11 \text{ MPa}$$

Equations (3) and (4) can be used for the calculation of 95% CI_{CE} and CI_{POP} for Confirmation experiments [35, 36].

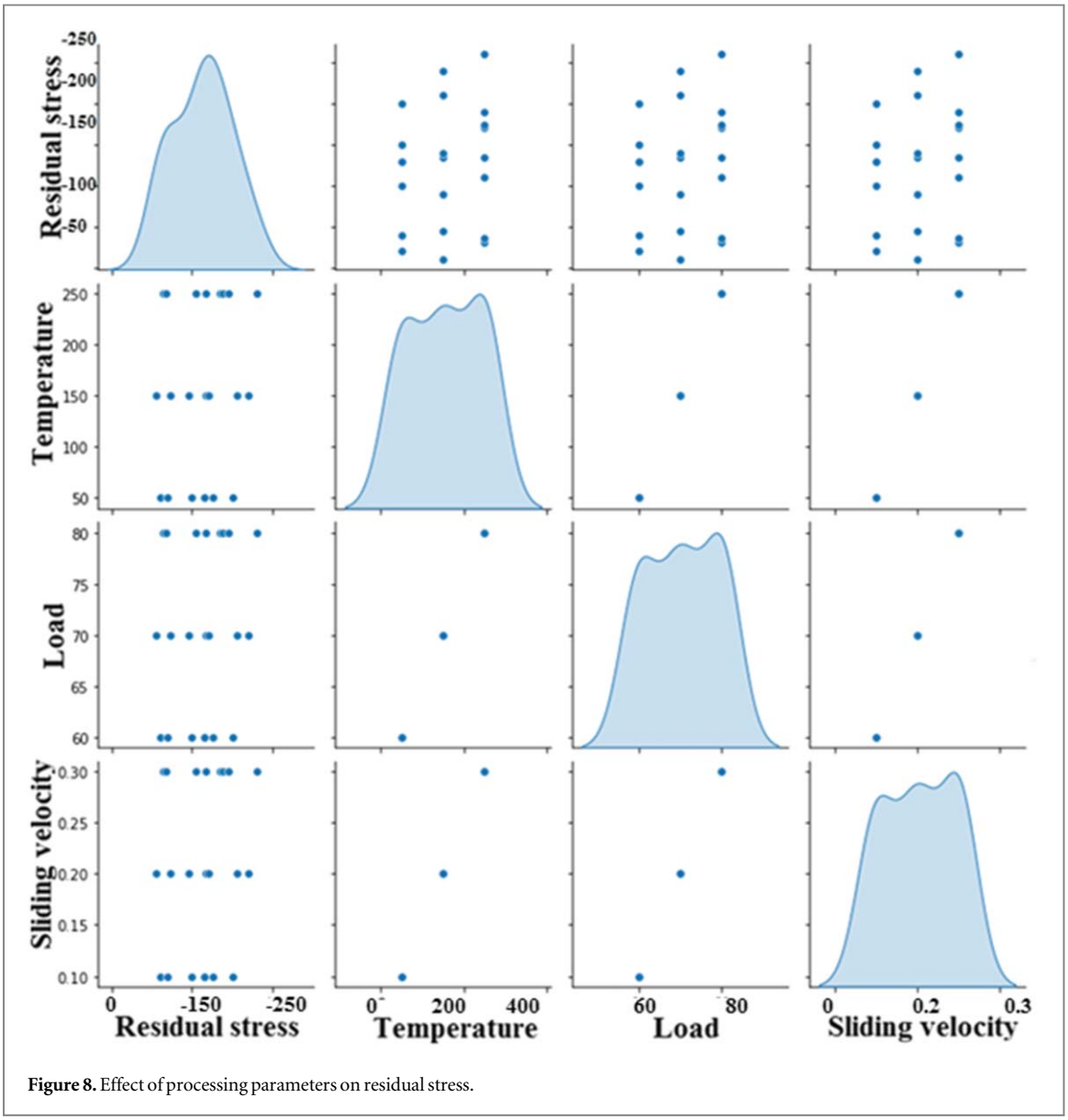


Figure 8. Effect of processing parameters on residual stress.

Table 4. Main Effects and average values of as-deposited carbon coating for residual stress .

Parameter	Level	Temperature		SV		Load	
Data		RD	S/N	RD	S/N	RD	S/N
Average values (weight loss)	L1	148.88	−42.95	140.00	−42.09	131.11	−41.64
	L2	114.44	−41.05	115.55	−41.08	108.88	−39.43
	L3	66.66	−36.00	74.44	−36.84	90.0	−38.93
The main effect of weight loss	L2-L1	−34.44	1.89	−24.44	1.00	−22.22	2.2
	L3-L2	−47.77	5.04	−41.11	4.23	−18.88	0.49
Difference (L3-L2)-(L2-L1)		−13.33	3.14	−16.66	3.22	3.33	−1.72

$$CI_{CE} = \sqrt{Fa(1, fe)Ve\left[\frac{1}{nef} + \frac{1}{R}\right]} \quad (16)$$

$$CI_{POP} = \sqrt{\left[\frac{1Fa\ (1, fe)Ve}{nef}\right]} \quad (17)$$

Following are the ANOVA calculation results:

$CI_{CE} = \pm 12.58$

While $CI_{POP} = \pm 8.32$

Predicted optimal values of as-deposited carbon coating for residual stress is:

CI_{CE}: $-1.58 < \mu < 23.58$

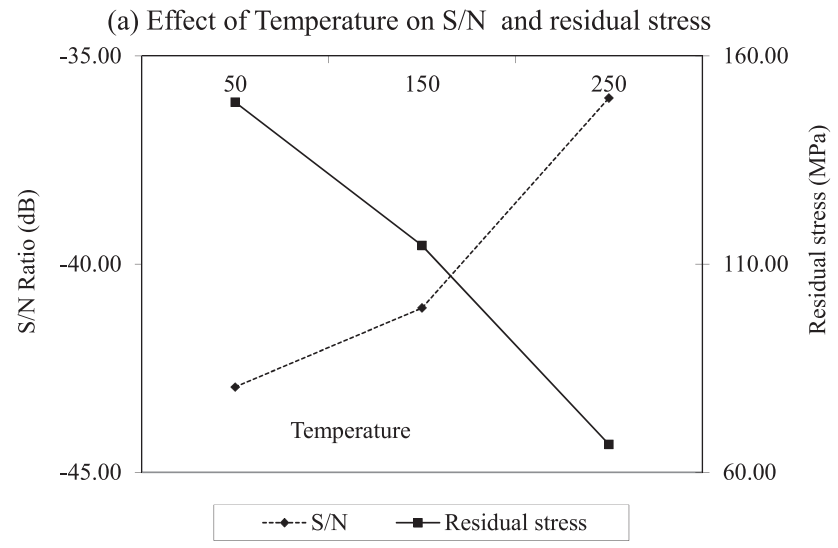
CI_{POP}: $2.68 < \mu < 19.32$

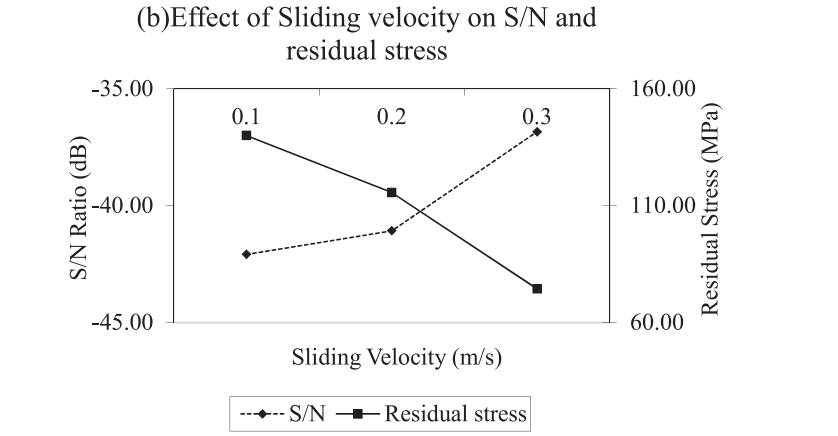
Predicted optimal values of as-deposited carbon coating for residual stress at significant parameters level were designated as:

Level 3 of A (Temperature) = 250 °C

Level 3 of B (Sliding velocity) = 0.3 m s^{-1}

Level 3 of C (Load) = 80 N





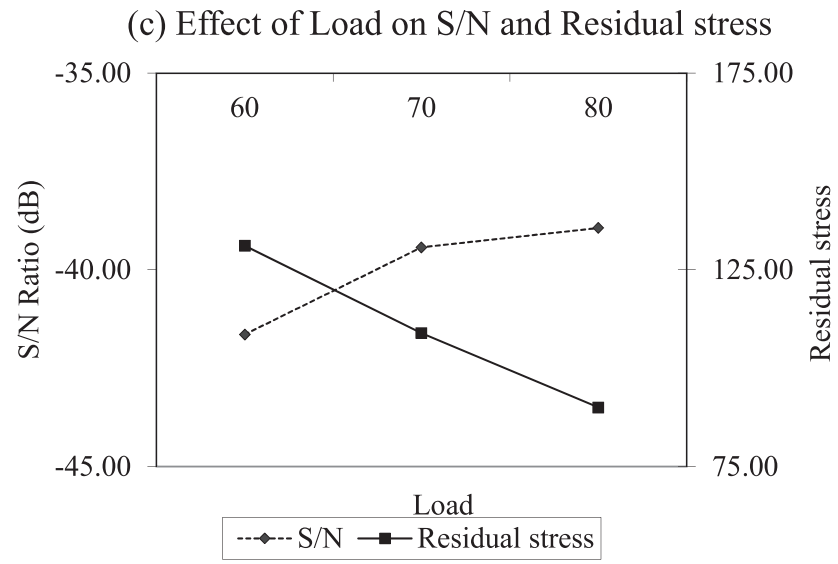


Figure 9. Shows deviation of residual stress with parameters (a) Temperature (°C) (b) sliding velocity (m s^{-1}) (c) Load (N).

3.3.3.2. Confirmation tests

Three confirmation tests of as-deposited carbon coating for residual stress after wear testing conditions were conducted for optimal level of parameters at level 3 of Temperature (250 °C), level 3 of sliding velocity (0.3 m s^{-1}), and level 3 of load (80 N). The average value of as-deposited carbon coating for residual stress after wear testing conditions was measured as 11 MPa which was in between the 95% confidence interval predicted optimum range for residual stress.

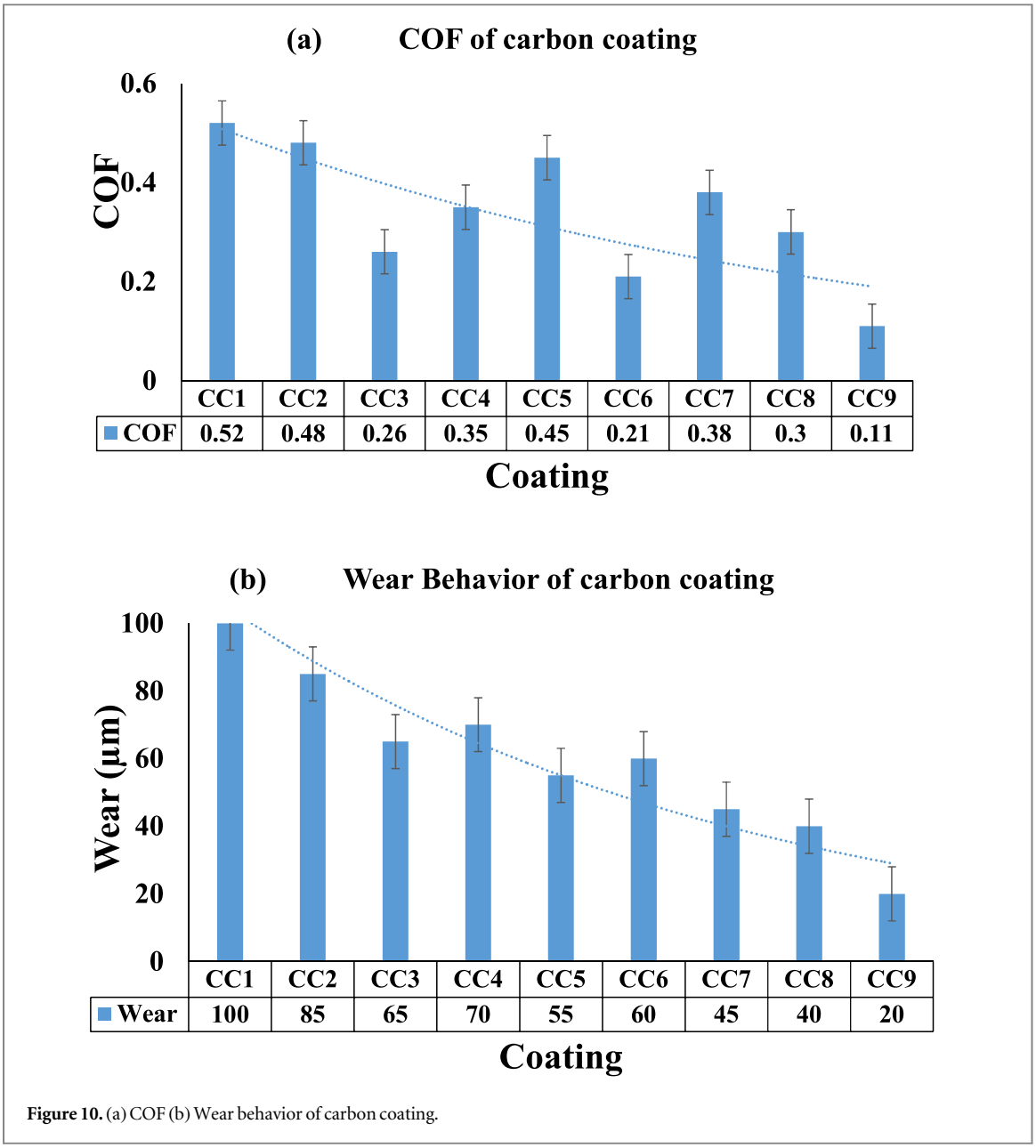


Figure 10. (a) COF (b) Wear behavior of carbon coating.

Table 5. Pooled ANOVA of as-deposited carbon coating for residual stress.

Source	SS		DOF		V		F-Ratio		SS′		P (%)	
	RD	S/N	RD	S/N	RD	S/N	RD	S/N	RD	S/N	RD	S/N
Temp.	30 688.8	77.27	2	2	15 344.4	38.63	200.14	129.54	30 535.5	76.68	51.23	56.04
SV	19 755.5	46.45	2	2	9877.7	23.22	128.84	77.86	19 602.2	45.85	32.89	33.51
Load	7622.2	12.48	2	2	3811.1	6.24	49.71	20.93	7468.8	11.89	12.53	8.69
E (Pooled)	1533.3	0.59	20	2	76.6	0.29	—	—	1993.3	2.38	3.34	1.74
Total	59 600.0	136.81	26	8	—	—	—	—	59 600	136.81	100	100

3.4. Tribological behavior of as deposited coating

Figures 10(a), (b) shows the tribological behavior of as-deposited carbon coating. The experimental study shows that COF varies from 0.11 to 0.52 while wear varies from 20 to 100 μm. The results concluded that as COF and wear is minimum (0.11 and 20 μm) and at 250 °C, 0.3 m s^{-1} and 70 N while maximum (0.52 and 100 μm) at 50 °C, 0.1 m s^{-1} and 60 N. Figures 11(a), (b) illustrates the FESEM images of hybrid composite coatings, clarifying the wear characteristics of damaged out surfaces. After conducting the wear test, FESEM images were obtained, indicating the brittle structures of grains besides the transformation of composite coatings. It was elucidated after conducting experiments that metal particles are embedded in the coating.

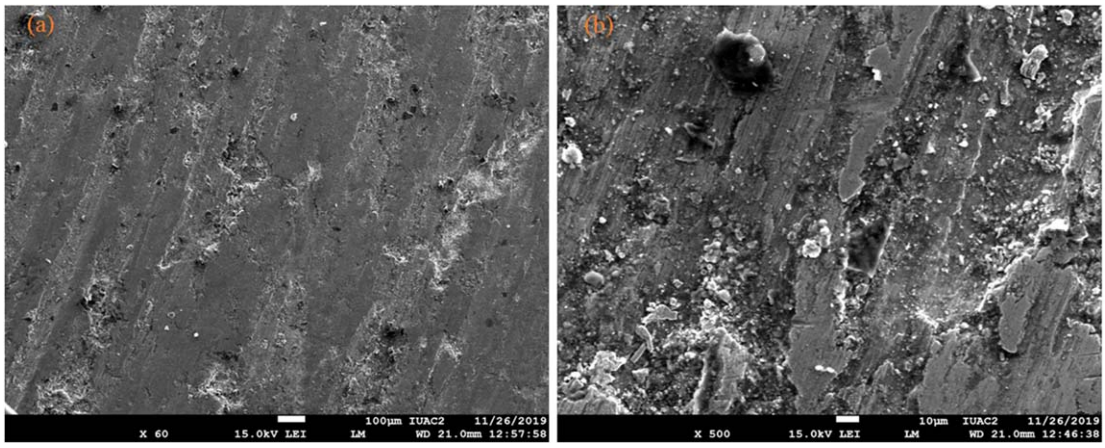


Figure 11. FESEM images of wear characteristics of coating.

Table 6. Residual stress average values at optimal levels.

Level	Residual stress (MPa)
A_3	66.6
B_3	74.4
C_3	90
T	110

4. Conclusions

In the present work, deposited carbon coating was successfully synthesized using HVOF thermal spray deposition technique. Following are the conclusions that can be computed as:

- The numerical modeling results confirm that HVOF thermal spray coating promotes compressive residual stresses due to high particle velocity and differences in thermal temperatures between coating and substrate materials.
- The estimated mean (μ_{RS}) for residual stress obtained from optimal responses is 11 MPa.
- Three confirmation tests of as-deposited carbon coating for residual stress after wear testing conditions were conducted for optimal level of parameters at level 3 of Temperature (250 °C), level 3 of sliding velocity (0.3 m s^{-1}), and level 3 of load (80 N).
- Percentage contribution o for residual stress after wear testing conditions was highest for Temperature (51.23%) followed by 32.8% sliding velocity and 12.5% load.
- The results concluded that as COF and wear is minimum (0.11 and 20 μm) and at 250 °C, 0.3 m s^{-1} and 70 N while maximum (0.52 and 100 μm) at 50 °C, 0.1 m s^{-1} and 50 N.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

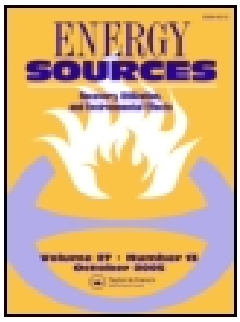
ORCID iDs

Ankit Tyagi https://orcid.org/0000-0001-9153-1402
Qasim Murtaza https://orcid.org/0000-0001-8784-2304
R S Walia https://orcid.org/0000-0002-1201-7892

References

[1] Deutsche Forschungsgemeinschaft (DFG) 1999 *Eigenspannungen und Verzug durch Warmeeinwirkung (Residual Stresses and Warpage Due to Heat Exposure)* (New York: Wiley) (in German)
[2] Clyne T W and Gill S C 1996 Residual stresses in thermal spray coatings and their effect on interfacial adhesion: a review of recent work *J. Therm. Spray Technol.* **5** 401–18
[3] Martina F, Roy M J, Scott B A, Terzi S, Colegrove P A, Williams S W, Withers P J, Meyer J and Hofmann M 2016 Residual stress of as-deposited and rolled wire + arc additive manufacturing Ti-6Al-4V components *Mater. Sci. Technol.* **32** 1439–48
[4] Piehl K H 1989 Formgebendes Schweissen von Schwerkomponenten (Formative Welding of Heavy-Duty Components) *Thyssen Tech. Ber.* **21** 53–71 (In German)
[5] Rai A K, Kumar A, Shankar V and Jayakumar T 2004 Characterization of microstructures in Inconel 625 using x-ray diffraction peak broadening and lattice parameter measurements *Scripta Mater.* **51** 59–63
[6] Al-Fadhli H Y, Stokes J, Hashmi M S J and Yilmaz B S 2006 HVOF coating of welded surfaces: fatigue and corrosion behaviour of stainless steel coated with Inconel-625 alloy *Surf. Coat. Technol.* **200** 4904–8
[7] Al-Fadhli H Y, Stokes J, Hashmi M S J and Yilmaz B S 2006 The erosion-corrosion behaviour of high-velocity oxy-fuel (HVOF) thermally sprayed Inconel-625 coatings on different metallic surfaces *Surf. Coat. Technol.* **200** 5782–8
[8] Raghupathy A C, Srinivasan D, Suwas S and Srivastava C 2017 Microstructural evolution of cold-sprayed inconel 625

superalloy coatings on low alloy steel substrate *Acta Mater.* **129** 11–25
[9] Ma X Q, Roth J, Gandy D W and Frederick G J 2006 A new high- velocity oxygen fuel process for making finely structured and highly bonded inconel alloy layers from liquid feedstock *J. Therm. Spray Technol.* **15** 670–5
[10] Azizpour M J and Nourouzi S 2013 Evaluation of surface residual stress in HVOF sprayed WC-12Co coatings by XRD and EDhole drilling *J. Mech. Sci. Technol.* **27** 2709–13
[11] Stokes J and Looney L 2001 HVOF system definition to maximise the thickness of formed components *Surf. Coat. Technol.* **148** 18–24
[12] Ahmed N, Bakare M S, McCartney D G and Voisey K T 2010 The effects of microstructural features on the performance gap in corrosion resistance between bulk and HVOF sprayed Inconel 625 *Surf. Coat. Technol.* **204** 2294–301
[13] Edris H and McCartney D G 1997 Microstructural characterization of high-velocity oxy-fuel sprayed coatings of Inconel 625 *J. Mater. Sci.* **32** 863–72
[14] Ahmed R, Hadfield M and Tobe S 2000 Variation in residual stress field during fatigue failure of thermal spray coating *Proc. of The Int. Thermal Spray Conf. (Montreal, QUE)* pp 399–406
[15] Tyagi A, Pandey S M, Gupta K, Walia R S, Murtaza Q and Krishen K 2019 Tribological behavior of sustainable carbon-based composite coating for wear resistance applications *Mater. Res. Express* **6** 125601
[16] Tyagi A, Walia R S and Murtaza Q 2021 Residual, corrosion & tribological behavior of HVOF sprayed sustainable temperature dependent carbon based hybrid composite coating *Strojniški vestnik - Journal of Mechanical Engineering* **674** 191–9
[17] Chandrasekaran M, Muralidhar M, Krishna C M and Dixit U S 2010 Application of soft computing techniques in machining performance prediction and optimization: a literature review *Int. J. Adv. Manuf. Technol.* **46** 445–64
[18] Praveen A S, Sarangan J, Suresh S and Channabasappa B H 2016 Optimization, and erosion wear response of NiCrSiB/WC–Co HVOF coating using Taguchi method *Ceram. Int.* **42** 1094–104
[19] Muthu Krishnan M, Maniraj J, Deepak R and Anganan K 2018 *Mater. Today Proc.* **5** 716–23
[20] Patel K A and Brahmbhatt P K 2016 *Proc. Technol.* **23** 391–7
[21] Tyagi A, Walia R S and Murtaza Q 2019 Tribological behavior of HVOF carbon coating for wear resistance applications *Mater. Res. Express* **6** 125606
[22] King P C, Bae G, Zahiri S H, Jahedi M and Lee C 2010 An experimental and finite element study of cold spray copper impact onto two aluminum substrates *J. Therm. Spray Technol.* **19** 620–34
[23] Bae G, Xiong Y, Kumar S, Kang K and Lee C 2008 General aspects of interface bonding in kinetic sprayed coatings *Acta Mater.* **56** 4858–68
[24] Ghelichi R, Bagherifard S, Guagliano M and Verani M 2011 Numerical simulation of cold spray coating *Surf. Coat. Technol.* **205** 5294–301
[25] Lyphout C, Nyle'n P and Wigren J 2007 Characterization of adhesion strength and residual stresses of HVOF sprayed Inconel 718 for aerospace repair applications *Proceedings of the 22nd International Conference on Surface Modification Technologies (SMT22) (Trollhättan, Sweden)* pp 11–18
[26] Tyagi A, Mani Pandey S, Walia R S and Murtaza Q 2019 Characterization and parametric optimization of tribological properties of Mo blend composite coating *Mater. Res. Express* **6** 086428
[27] Tyagi A, Pandey S M, Murtaza Q, Walia R S and Tyagi M 2020 Tribological behavior of carbon coating for piston ring applications using Taguchi approach *Materials Today Proceeding* **25** 759–64
[28] Prakash K S, Thankachan T and Radhakrishnan R 2017 Parametric optimization of dry sliding wear loss of coppe—MWCNT composites *Trans. Nonferrous Met. Soc. China* **27** 627–37
[29] Cevik A and Guzelbey I H 2008 Neural network modeling of strength enhancement for CFRP confined concrete cylinders *Build. Environ.* **43** 751–63
[30] Kurt H I 2016 Influence of hybrid ratio and friction stir processing parameters on ultimate tensile strength of 5083 aluminum matrix hybrid composites *Compos. Part* B **93** 26–34
[31] Skordaris G, Bouzakis K-D, Charalampous P, Kotsanis T, Bouzakis E and Lemmer O 2016 Effect of structure and residual stresses of diamond-coated cemented carbide tools on the film adhesion and developed wear mechanisms in milling *CIRP Annals—Manufacturing Technology* **65** 101–4
[32] Tyagi A, Walia R S and Murtaza Q 2019 Tribological behavior of temperature-dependent environment-friendly thermal CVD diamond coating *Diamonds and Related Materials* **96** 148–59
[33] Bouzakis K-D, Skordaris G, Bouzakis E, Charalampous P, Kotsanis T, Tasoulas D, Kombogiannis S and Lemmer O 2015 Effect of the interface fatigue strength of NCD coated hardmetal inserts on their cutting performance in milling *Diamond, and Related Materials* **59** 80–9
[34] Tyagi A, Pandey S M, Walia R S and Murtaza Q 2019 Characterization and parametric optimization of % change in residual stress of Mo composite coating using Taguchi approach *Mater. Res. Express* **6** 125623
[35] Goyal T, Walia R S and Sidhu T S 2011 Study of coating thickness of cold spray process using Taguchi method *Mater. Manuf. Processes* **27** 185–92
[36] Walia R S, Shan H S and Kumar P 2005 Parametric optimization of centrifugal force-assisted abrasive flow machining (CFAAFM) by the Taguchi method *Mater. Manuf. Processes* **21** 375–82



Taylor & Francis
Taylor & Francis Group

Energy Sources, Part A: Recovery, Utilization, and Environmental Effects

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ueso20

Experimental investigation and artificial neural network modeling of performance and emission of a CI engine using orange peel oil- diesel blends

Sandeep Kumar Barnawal & Naveen Kumar

To cite this article: Sandeep Kumar Barnawal & Naveen Kumar (2021): Experimental investigation and artificial neural network modeling of performance and emission of a CI engine using orange peel oil- diesel blends, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, DOI: 10.1080/15567036.2021.1967520

To link to this article: https://doi.org/10.1080/15567036.2021.1967520

Published online: 21 Aug 2021.

Submit your article to this journal

Article views: 3

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=ueso20

ENERGY SOURCES, PART A: RECOVERY, UTILIZATION, AND ENVIRONMENTAL EFFECTS
https://doi.org/10.1080/15567036.2021.1967520

Taylor & Francis
Taylor & Francis Group

Check for updates

Experimental investigation and artificial neural network modeling of performance and emission of a CI engine using orange peel oil-diesel blends

Sandeep Kumar Barnawal and Naveen Kumar

Mechanical Engineering Department, Centre for Advanced Studies and Research in Automotive Engineering, Delhi Technological University, New Delhi, India

ABSTRACT

The exponential decay of crude oil and environmental degradation forced researchers to investigate alternative sustainable fuels. One of the promising alternatives to petroleum oil is orange peel oil. Orange peel oil has characteristics that are similar to diesel, and it is also simple to blend. In this paper, performance and emission parameters are investigated, and using an artificial neural network (ANN), its prediction and validation have been made. The blend shows the best results for 30% orange peel oil by vol. and 70% diesel. At peak load, the brake thermal efficiency (BTE) of this blend is 17.5% higher than diesel, and the brake-specific energy consumption (BSEC) is 20% lower. These experimental results were used to perform prediction and validation. The prediction of data is made employing the Quasi- newton method algorithm, and it is found that it best fits the linear regression analysis. The value of R^2 for BTE and BSEC is 0.994 and 0.986, respectively. The suggested ANN model performance and accuracy were totally satisfactory. There is a noticeable reduction in the emission parameters.

ARTICLE HISTORY
Received 13 May 2021
Revised 2 August 2021
Accepted 3 August 2021

KEYWORDS
Artificial neural network; orange peel oil; quasi-newton method; performance; emissions

Introduction

With the exponential rise in vehicles and industry, the energy demand is increasing day by day. The depletion of energy resources in the last few decades brings researcher's attention to explore alternative fuels. However, crude oils account for a significant portion of the transportation and agriculture sectors. (Tomar et al. 2020). Carbon dioxide and greenhouse gases are produced mainly by petroleum fuels. It was discovered that CO2 emissions would more than quadruple by 2040, with a total output of 5Gt. Co2 emissions are expected to rise from 7% to 14% by 2040. The global temperature in the last decades increased by 2°C. To clean the environment and reduce this dependency, a sustainable substitute for diesel and bio-fuel, hydrogen fuel, etc., can be a promising alternative fuel (Brindhadevi et al. 2021). Organic wastes are now one of agriculture's major concerns from an environmental viewpoint; we may be able to use this waste/biomass for energy generation. (Martín et al. 2018). Orange peel oil has great potential and may be utilized as a replacement fuel at a low cost of acquisition. (Deep et al. 2013). In 2019, the globe produced 79 million tonnes of oranges, with Brazil providing 22% of the total, followed by China and India. India has around 10% of the total world sweet orange production. Oranges are the world's second most popular fruit, followed by grapes, and the third-largest fruit cultivated in India. (Citrus : World Markets and Trade 2021). In India, oranges are majorly produced in Maharashtra, Punjab, Andhra Pradesh, and Karnataka (Deep et al. 2015). The physical and chemical characteristics of vegetable oil influence the air-fuel blending procedure.

CONTACT Sandeep Kumar Barnawal sandeep_mt2k19@dtu.ac.in Mechanical Engineering Department, Centre for Advanced Studies and Research in Automotive Engineering, Delhi Technological University, New Delhi 110042, India

© 2021 Taylor & Francis Group, LLC

Furthermore, the spray properties of the engine are affected by the fuel's density, vaporization, and viscosity (Manigandan et al. 2021). The comprehensive literature studies were carried out on Orange peel oil and other vegetable oils. Its brief analysis of some of the research conclusions are summarized below:

(Kumar, Kannan, and Nataraj 2020) The performance, combustion, and emission characteristics of a diesel engine fueled by waste orange peel oil biodiesel nanoemulsion were investigated. When related to pure OOME at peak load, the BTE of the engine improved by 1.6% and 3.0% for OOME-T50 and OOME-T100 fuels, respectively. In addition, there will be a significant drop in C.O. and H. C. emissions.

(Purushothaman and Nagarajan 2009) The results of experimental research on a CI engine using opo and opo with DEE show that it may be used as a diesel replacement. The BTE for DEE with opo is around 3% more than that of diesel fuel, and for opo-diesel blends, it is about 1.2% higher. Compared to diesel fuel, NOx emissions are lower for DEE with opo and higher for the opo-diesel blend. Furthermore, smoke emissions are lower for DEE with opo and opo-diesel fuel blends, whereas H. C. emissions are more significant for DEE with opo and lower for opo-diesel mix.

(Hoekman and Robbins 2012) has looked at the "biodiesel NOx consequence" and theories to explain it. The NOx effect of biodiesel may be reduced by adjusting engine controller settings, primarily by delaying injection timing and boosting exhaust gas recirculation (EGR). With newer engines, the total magnitude of the biodiesel NOx result looks to be reduced. However, there are still situations when the percentage change is still significant.

According to literature, when compared to diesel, biodiesel appears to be a viable alternative for diesel engine applications in engine performance and engine exhaust gas emissions. Straight opo is mixed with diesel and tested on a single cylinder CI engine in this study. Due to poor flow characteristics (viscosity, density, and calorific value), straight vegetable oils (SVO) as an alternative fuel are forbidden, resulting in engine inefficiencies. However, the current study looks at using natural vegetable oil, such as orange peel, as a potential diesel fuel substitute. The engine performance and emissions were investigated using a single-cylinder direct ignition diesel engine in a series of engine experiments under various loading circumstances, demonstrating the credibility of orange peel oil blends as a potential alternative and also use of ANN to validate its experimental results with predicted results.

Materials and methods

In this study, diesel and orange-peel oil is blended in different ratios (by vol./vol. %) to recognize the effect of blends on engine performance and emission characteristics. In the following sections, the preparation of fuel and different blend properties are explained.

Preparation of fuel

Steam distillation is used to extract oil from orange peels. 4-Isopropeny-1-methyl-cyclohexene, commonly known as D-limonene, is extracted. To begin the distillation process, prepare a mixture of orange peel and distilled water. The mixture was then funneled into the flask and well stirred. The distillation apparatus and the ring and wire gauze were then installed on top of the ring stand. A burner was installed beneath the instrument to produce heat. Distilled water was used to fill the separating funnel. Aluminum foil was used to coat the still head and Claisen head. It takes 18–22 minutes to reach boiling point, depending on the quality of the orange peel blend and the flame strength. As the mix reaches boiling point, the still pot's water content decreases; therefore, a small amount of water is applied through the separatory funnel. The distilled substance was then removed, and the distillation was treated with NaCl. During the distillation, a thin layer of limonene was discovered. The yield of this method is 1.5 kg of peel that yields 700 ml of peel oil.

Table 1. Different blends nomenclature.

Blend name	Orange peel oil (%)	Diesel
opo10	10	90
opo20	20	80
opo30	30	70
opo40	40	60
D100	0	100

Table 2. Physicochemical properties of orange peel oil blends.

Properties	ASTM std.	D100	opo100	opo10	opo20	opo30	opo40	Measuring Instrument
Density@15°C (Kg/m^3)	ASTM D4052	0.853	0.861	0.821	0.824	0.832	0.838	Anton Par, DMA 4500
Kinematic viscosity at 40°C (cST)	ASTM D445	2.48	3.19	2.62	2.95	3.12	3.14	Visco bath, Petrotest
Lower calorific value (MJ/Kg)	ASTM D240	43.12	41.54	43.75	44.02	44.03	44.64	Oxygen Bomb Calorimeter Parr

Blend development of orange peels oil and diesel

Diesel and orange-peel oil's blending process is easy and can be blended easily without any phase separation and color change problem. The blends were prepared by volume percentage. Table 1 shows the nomenclature of various blends.

Physiochemical properties of fuel sample

Normal test expediencies were projected for orange peel oil and its blend with diesel. The calorific value for the straight opo is low compared to the diesel but has high density and viscosity. The diverse blends were opo10, opo20, opo30, opo40, and D100. As the percentage of opo increases, density and Kinematic viscosity also increase, but the calorific value decreases. The physicochemical properties of opo blends are listed in the Table 2

Experimental test rig

A 4-stroke single cylinder direct injection engine was chosen for the engine trials. It has a rated power output of 3.5 kW and a constant rpm of 1500. The engine's complete specification is shown in Table 3. The electrical loading unit is linked to the engine and generator. An electronic information acquisition system is used on the test rig engine to monitor and measure the fuel flow rate, engine load, and temperature under various conditions. The ammeter, voltmeter, and rpm indicator are fitted on the electrical loading

Table 3. Specifications of the engine.

Parameters	Specifications
Type	The single-cylinder, four-stroke, direct-injection engine
Model	KSA-2057-*P*-1170 (Kirloskar)
Cubic capacity	0.78 ltr
Bore diameter	95 mm
Orifice diameter	13.4 mm
Rated power	3.5 kW
Rated speed	1500 rpm
Compression ratio	17.5:1
Fuel injection timing	23° bTDC (before top dead center)
Fuel injection pressure	200 bar
Lubrication System	Forced feed
Cooling system	Air-cooled (Radial Cooled)

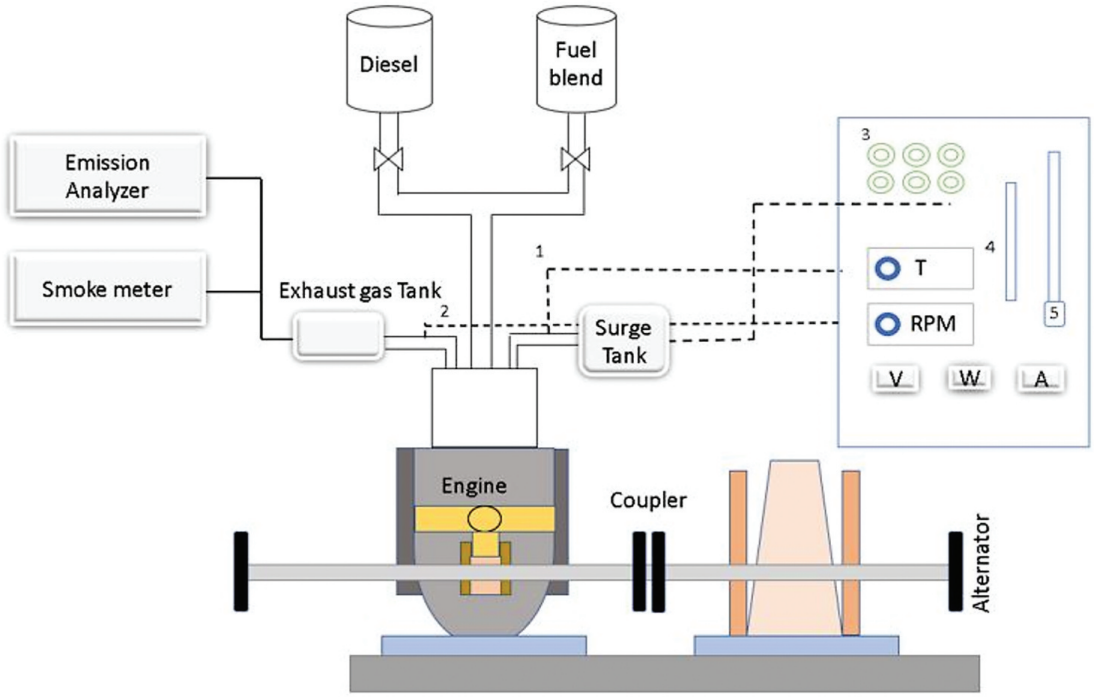


Figure 1. Engine setup.

Table 4. Measuring principle, accuracy, range and percentage uncertainty of different measurements.

Measurements	Measurement Principle	Range	Accuracy	Uncertainty (%)
Engine load	Strain gauge type load cell	0–25 kg	±0.1 kg	±0.1
Time	Digital stop watch	-	±0.2 s	±0.1
Engine speed	Magnetic pick-up type	0–2000 rpm	±10 rpm	±0.3
Temperature	K-type thermocouple	0–1000°C	±0.5 °C	±0.4
Fuel flow	Differential pressure transmitter	0–500 mm	±0.05 mm	±0.02
NO	Chemi-luminescence principle, electrochemical sensor	0–5000 ppm	±50 ppm	±0.5
HC	Flame ionization detector-FID	0–20000 ppm	±10 ppm	±0.2
Smoke density	Hatridge principle	0–100%	±1%	±0.5
Fuel consumption	Level sensor	-		±0.2
Air consumption	Turbine flow type	-		±0.1

unit to measure the engine's current, voltage, and speed. The setup of the engine is represented in Figure (1) . A thermocouple is installed on the exhaust pipe to monitor the temperature of the exhaust gas. An AVL smoke meter and gas analyzer are placed at the exhaust pipe to monitor the exhaust gas emissions.

Performance and exhaust emission error and uncertainty analysis

It is not unusual for errors to arise throughout experimental research. The comprehensive study of the uncertainties during the physical measurement is known as error analysis. During the experiment, human intervention, erroneous data input, out-of-date instrument calibration, and adverse climatic conditions were all prevalent mistakes. The accuracy and ambiguity of the different instruments employed are addressed in depth in Table 4

The overall uncertainty of the current investigational study:

=Square root of [(uncertainty of BTE)^2+ (uncertainty of Brake Power)^2+(uncertainty of BSEC)^2 (uncertainty of NO)^2+(uncertainty of CO)^2+(uncertainty of smoke)^2+(uncertainty of HC)^2]

= Square root of [$(0.25)^2$+$(0.05)^2$+$(0.3)^2$+$(0.5)^2$+$(0.1)^2$+$(0.5)^2$+$(0.2)^2$]

=0.8396

Performance characteristics

Brake thermal efficiency

Figure (2) shows the deviation of thermal brake efficiency (BTE) with Brake mean effective pressure (BMEP) of the various fuel samples. BTE is the amount of chemical energy of fuel that gets converted into valuable energy, i.e., shaft work. All fuel samples show an increasing trend in BTE with the rise in engine load. This can be due to the richer fuel supply at the higher load. Richer fuel has more chemical energy, which releases high power, but the BTE decreases for all the fuel samples at the higher load; this can be because of an insufficient amount of oxygen inside the cylinder. A similar kind of trend is mentioned by various researchers. (Dharma et al. 2017). Moreover, with the increase of opo percentage, the BTE increases significantly. At peak load conditions, the BTE of opo10, opo20, opo30, and opo40 is higher than diesel by the percentage of 3.51, 9.9, 17.5, and 13.06. Due to the higher Calorific value in the fuel, opo30 has better BTE than other fuel samples.

Brake-specific energy consumption

The amount of energy consumed for the development of unit power is represented by Brake specific energy consumption (BSEC). As it considers both heating value and mass flow rate of fuel, so it is an essential parameter of engine performance. The effect of the fuel sample on BSEC is shown in Figure (3) . BSEC at lower load is high, and as the load increases, the BSEC starts decreasing; this can be seen for all the fuel samples and this might be due to poor fuel mixing at minor loads, which causes partial combustion and raises the BSEC. opo30 has the lowest BSEC due to the mutual effect of high energy content and oxygen content. Several researchers have noticed a similar trend. (Deep et al. 2013). The viscosity of opo30 is higher and cannot be neglected, but its oxygen content and heating value overcome it. At peak load conditions, the BSEC of opo10, opo20, opo30, and opo40 is lower than diesel by the percentage of 8.81, 15.55, 20, and 3.81 respectively.

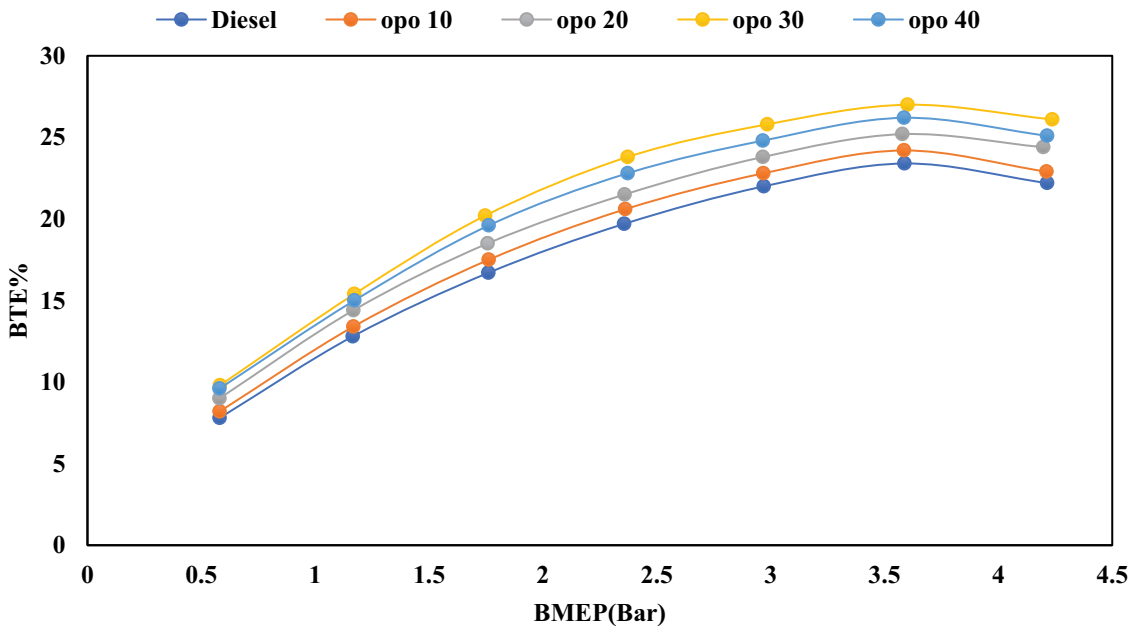


Figure 2. Variation of BTE with respect to BMEP.

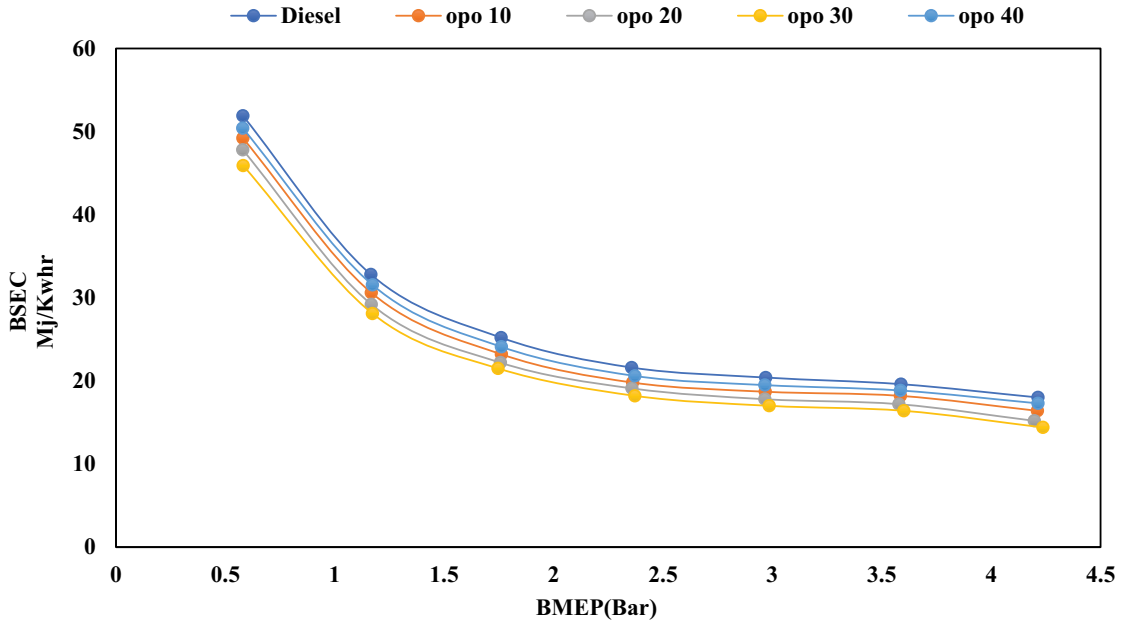


Figure 3. Variation of BSEC with respect to BMEP.

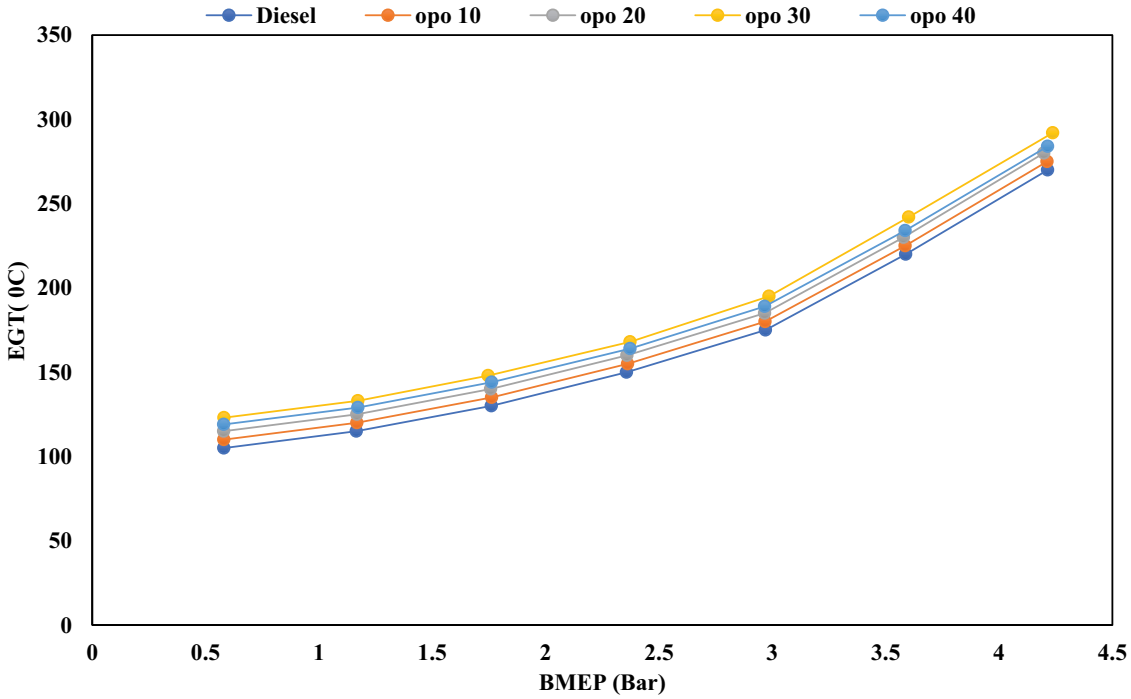


Figure 4. Variation of EGT with respect to BMEP.

Exhaust gas temperature

Variation of Exhaust gas temperature (EGT) with BMEP is highlighted in Figure (4) for different fuel samples. The graph shows that the increase of load EGT also increases for all the blend samples. The reason for the increase in EGT is heat released during diffusion combustion. This heat cannot be utilized for helpful power production, termed as waste heat. For all the engine loads diesel shows the least EGT, and opo30 offers the highest EGT for all loading conditions. opo30 has a higher density, calorific value, and oxygen concentration than the other blended fuel samples, resulting in a highest

exhaust gas temperature. Several authors have noticed a similar pattern. (Tomar et al. 2020). opo30 has the highest EGT; this may be because of the low high oxygen contents, which provide better fuel combustion. At the peak load, the EGT of Diesel is 6.2% lower than opo30.

Emission characteristics

Nitrogen oxides

Figure (5) shows the variation of NOx emission with BMEP for various blend samples. The NOx for all samples shows an increasing trend with the load. NOx emission depends on the ignition temperature, reaction time, and oxygen availability. As combustion temperature increases, the NOx emission also increases. The NOx emission for diesel is lower than the other opo fuel samples; this may be due to the lesser EGT of Diesel, and since opo fuel samples have higher oxygen content than diesel, better combustion takes place in opo samples. It increases cylinder temperature and results in higher NOx. Several studies have suggested a similar type of tendency (Siva, Munuswamy, and Devarajan 2019). Diesel has the most negligible NOx emission; this may be because of the low EGT and low oxygen content. At the peak load, the NOx of opo30 is 4.17% higher than diesel.

Opacity

Indirect measurement of the soot particles in emission is smoke opacity. The formation of smoke mainly depends upon the air-fuel ratio and fuel consumption. Figure (6) shows the variation of smoke for different blend samples with BMEP. The graph is of increasing in nature for all samples and all load conditions. Among the fuel sample for D100, the smoke is maximum. This is because of the low oxygen content and poor detonation characteristics of pure diesel. On the other side, because of the higher oxygen content in biofuels, the burning is efficient, leading to low smoke opacity. Several

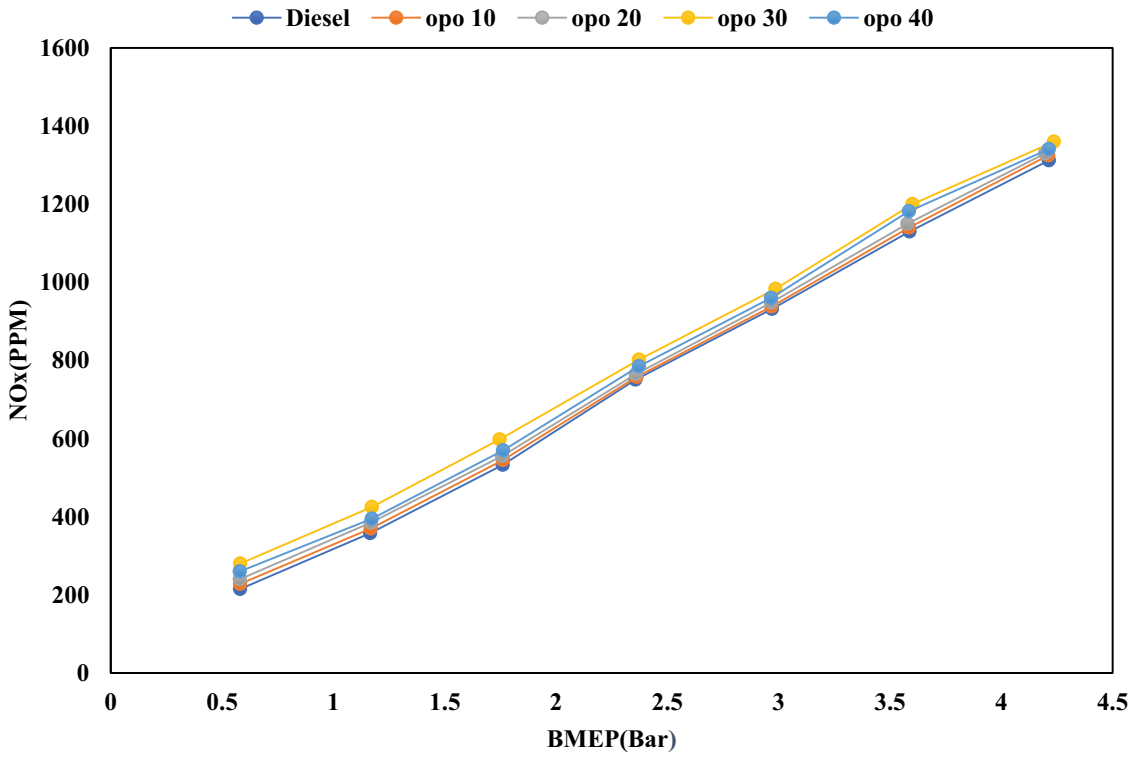


Figure 5. Variation of NOx with respect to BMEP.

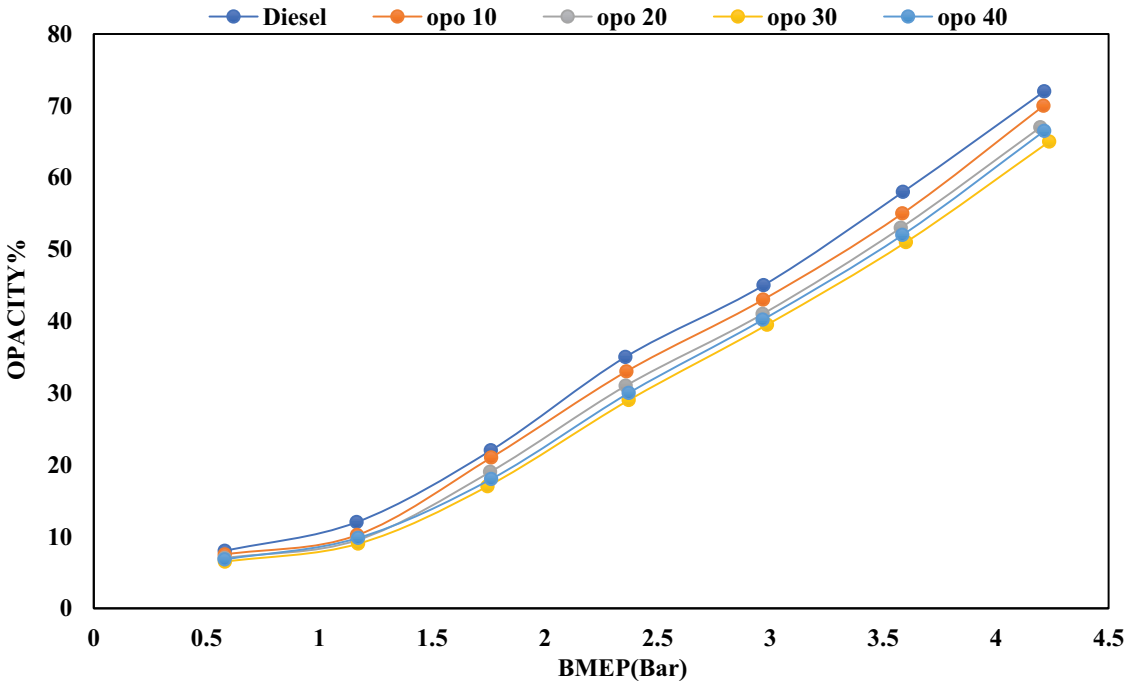


Figure 6. Variation of Opacity with respect to BMEP.

academics have pointed out a similar pattern.(Kumar, Kannan, and Nataraj 2020). In a complete load condition, the Opacity of opo10, opo20, opo30, and opo40 drops by the percent of 2.7, 6.9, 9.7, and 8.3, respectively.

Carbon monoxide

Figure (7) shows the variation of the CO emission for various blend samples with BMEP. The emission is more petite at a lower load, and as the load increases, the emission level also increases. CO is produced as a result of partial combustion of fuel. The CO emission is higher; it may be due to

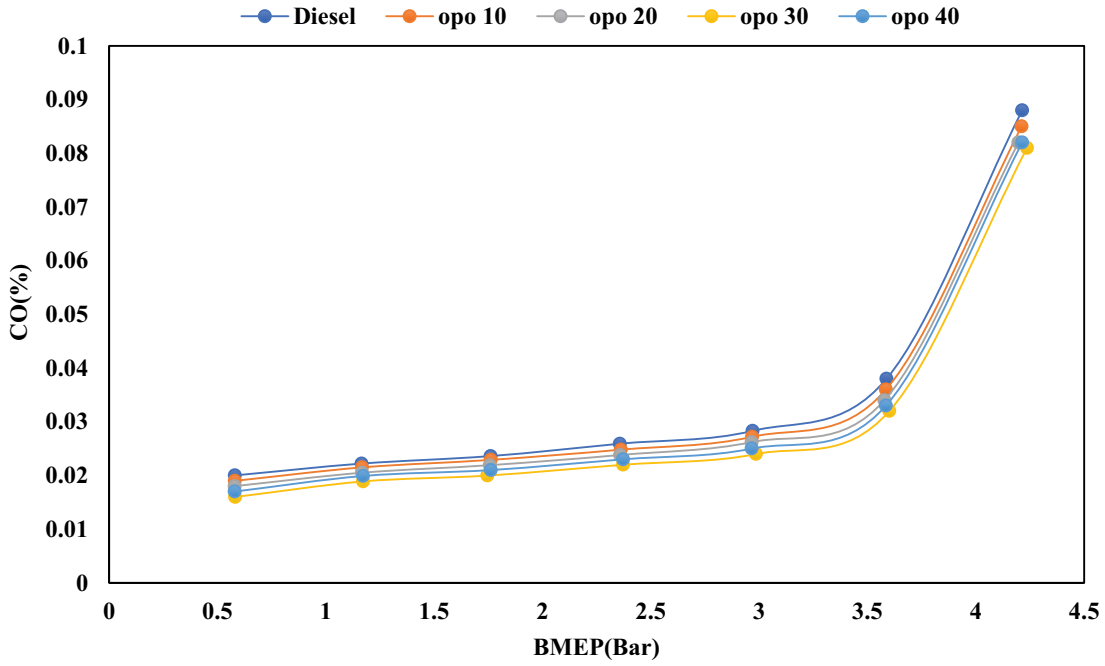


Figure 7. Variation of C.O. with respect to BMEP.

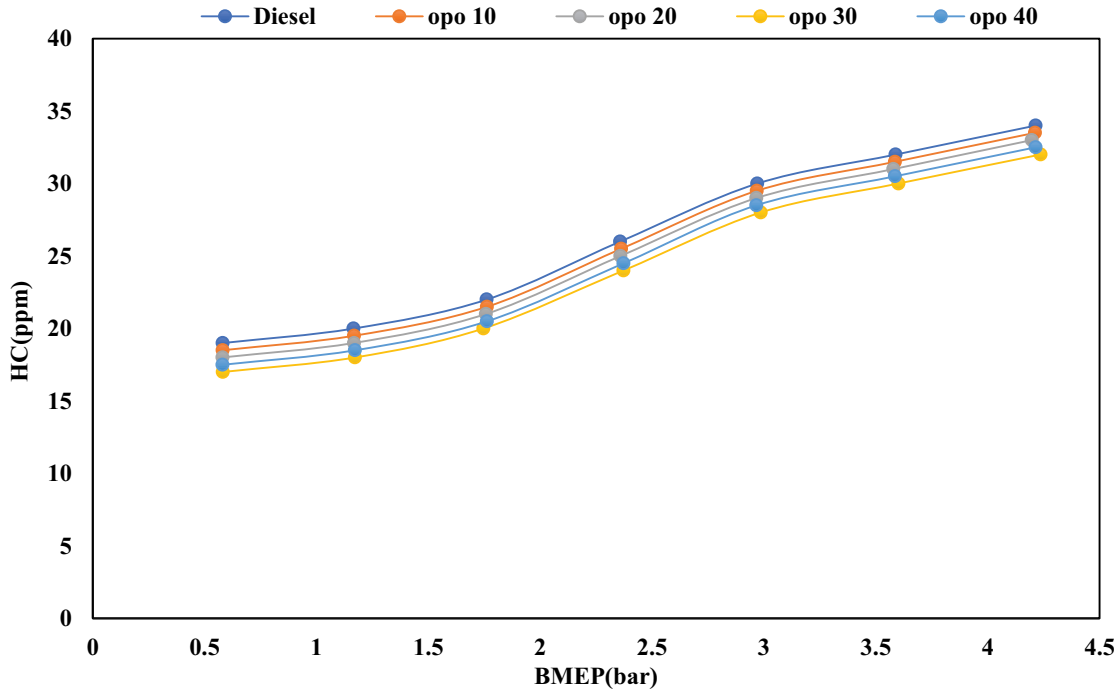


Figure 8. Variation of H.C. with respect to BMEP.

improper fuel and air mixing. CO emission is found significantly higher at the peak load. This may be due to the deficiency of oxygen at peak load. Various researchers mention a similar kind of trend. (Deep et al. 2013). The value of CO emission at the peak load for D100, opo10, opo20, opo30, and opo40 is 72, 70, 67, 65, and 66.5, respectively. Among the various samples, emission for opo30 is minimum; this may be because of higher oxygen content and proper combustion.

Hydrocarbon

Emission of hydrocarbon is the consequence of partial combustion of fuel samples. Figure (8) shows the variation of the hydrocarbon with the Brake mean effective pressure. As the load increases, the emission of hydrocarbon also increases. At a lower load, the hydrocarbon emission is low; this may be

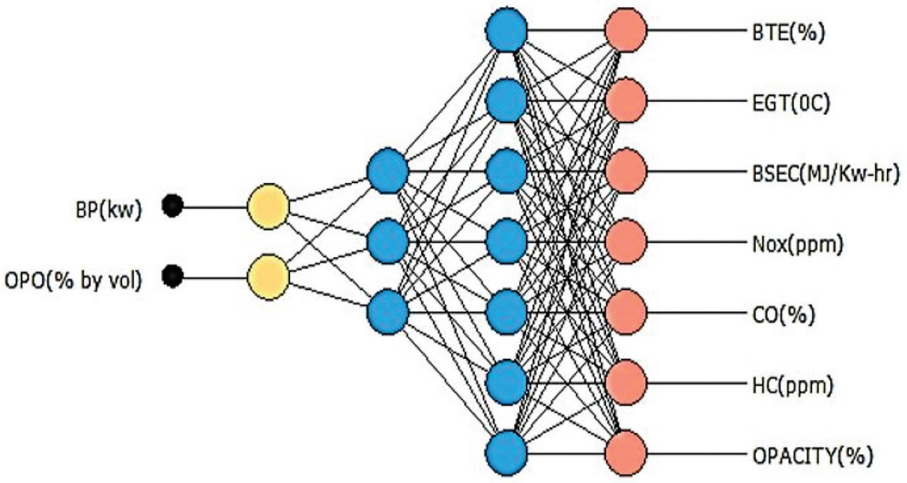


Figure 9. ANN model structure.

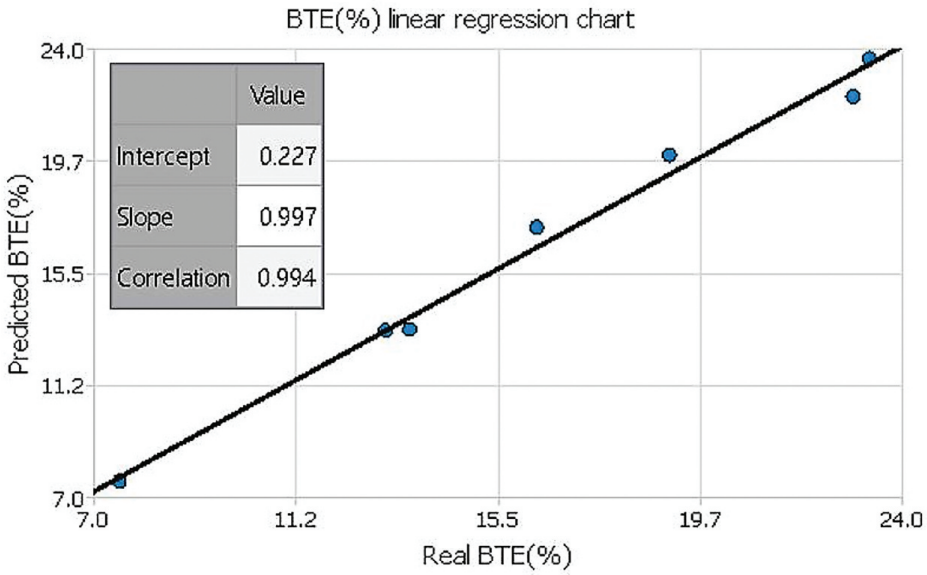


Figure 10. Real BTE (%) vs. predicted BTE (%).

due to the appropriate mixing of fuel and air. When the load is low, the combustion time becomes more and more oxygen-rich. Combustion time is shorter with greater loads, and oxygen concentration is lower, resulting in increased hydrocarbon emissions. Several studies have suggested a similar type of tendency (Purushothaman and Nagarajan 2009). Among the various fuel samples, diesel has higher HC emissions compare to other fuel samples. The opo30 has minor HC emissions; this may be due to high energy content value, high oxygen content, and high density. At the peak load condition, the opo30 has 5.8% less emission than the diesel.

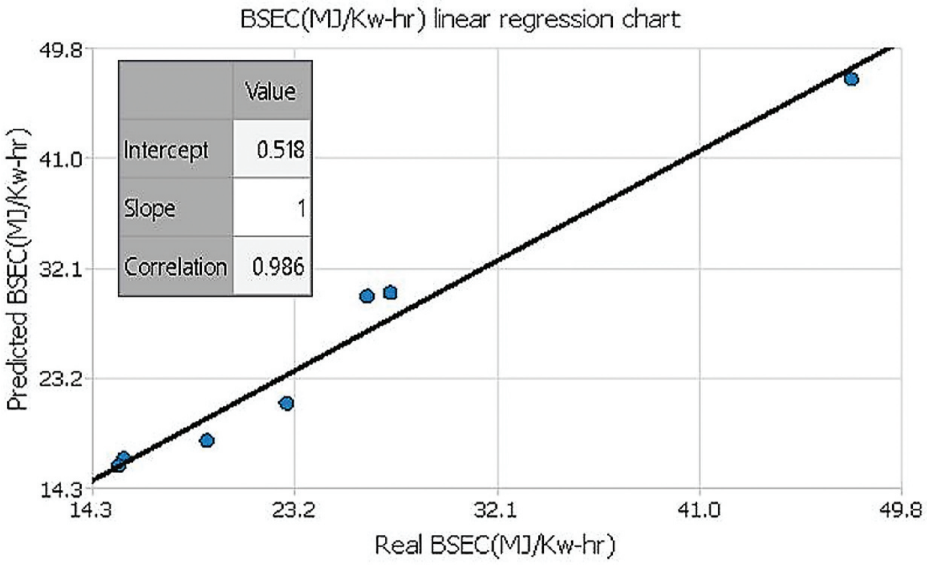


Figure 11. Real BSEC(MJ/kW-hr) vs.predicted BSEC(MJ/kW-hr).

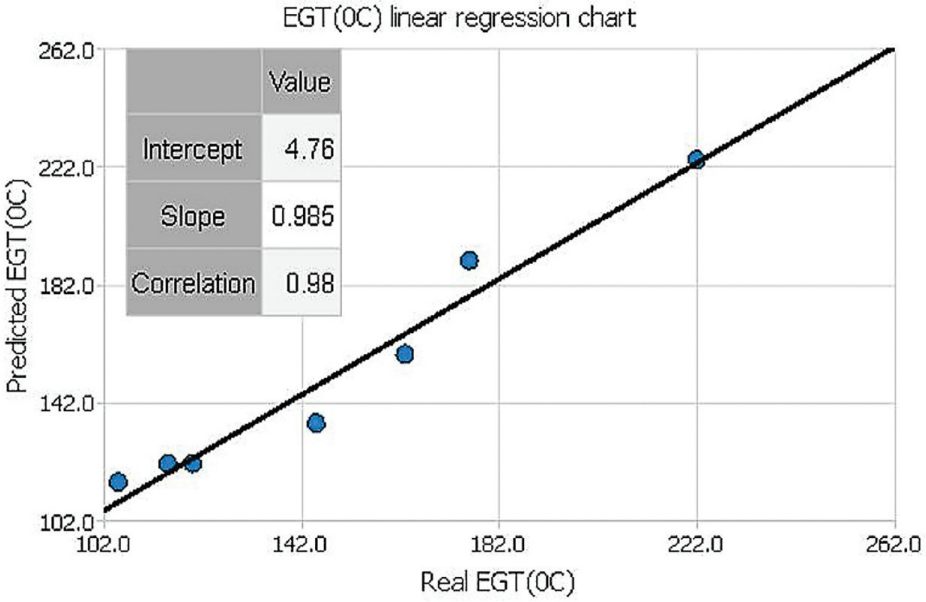


Figure 12. Real EGT vs. predicted EGT.

Artificial Neural Network (ANN) application

ANN is the piece of the computing system that is designed in such a way that it can simulate information and data that exhibit human-like intelligence. It is the basis of artificial intelligence (A. I.) and resolves glitches that would prove impossible or challenging by human or statistical standards. Artificial Neural Network (ANN) uses the brain's processing as a foundation to grow algorithms that can be used to model multifaceted outlines and forecast glitches. It mainly comprises three layers; input layers, hidden layers, and output layers(Uslu and Celik 2018). Figure (9) shows the input, hidden and output layer. Each of the three layers has a certain number of insignificant creatures known as neurons and roughly attentive processing components. The neurons are strengthened by communication connections, which are connected to the association weights. (Hosseini et al. 2020). There are

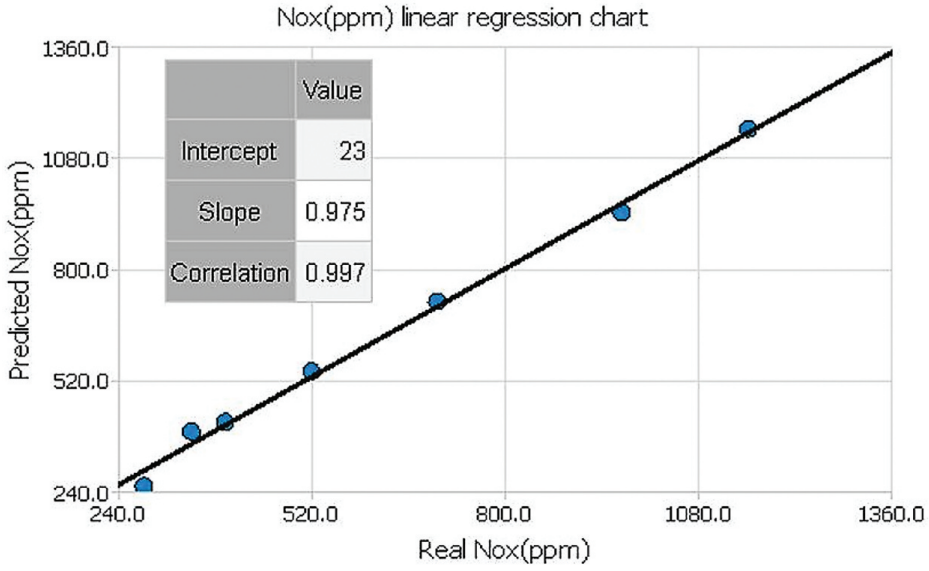


Figure 13. Real NO_x(ppm) vs. predicted NO_x(ppm).

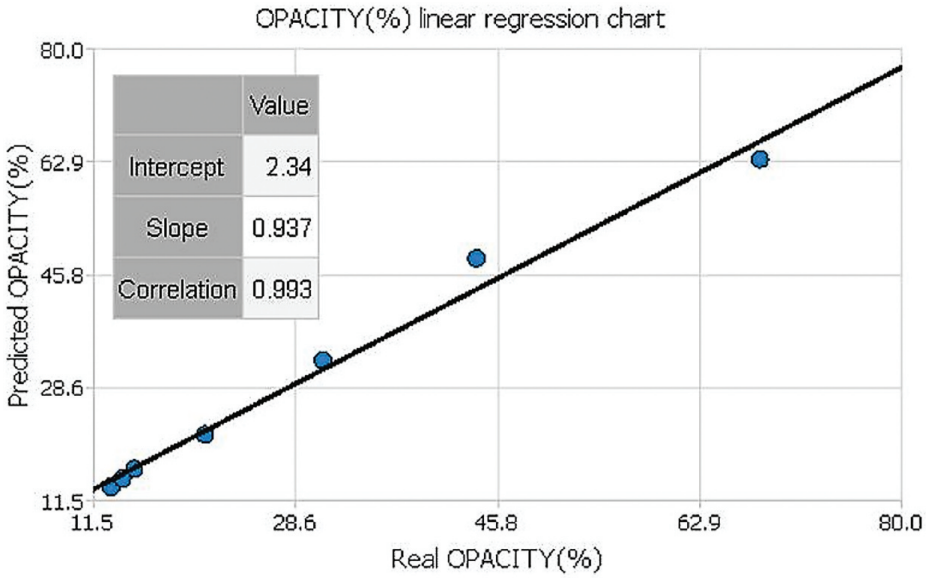


Figure 14. Real Opacity (%) vs predicted Opacity (%).

two stages in all machine learning algorithms and the ANN life cycle. One is the training stage, and the other is the prediction stage(Dharma et al. 2017). in the training stage, weight and bias of data are provided. In the training stage, the network is trained for output according to input data.

The ideal algorithm is the "Quasi-Newton method" in ANN model development and optimization. Quasi-Newton methods are used to find either zeros or local maxima and minima of the functions. In this employed algorithm, ANN is cast-off to evaluate the output parameters (BTE, BSEC, EGT, NOx, H.C., CO, and Opacity) based upon the input parameters (opo% and engine load). In the designed model, 60% of data is used for training instances, 20% of data is used for the selected instance, and the

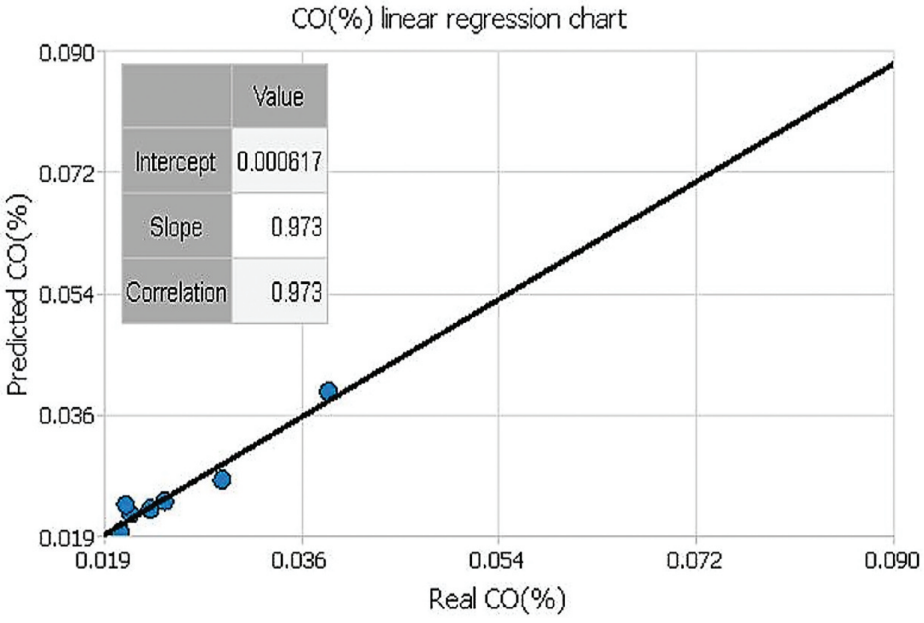


Figure 15. Real CO (%) vs predicted CO (%).

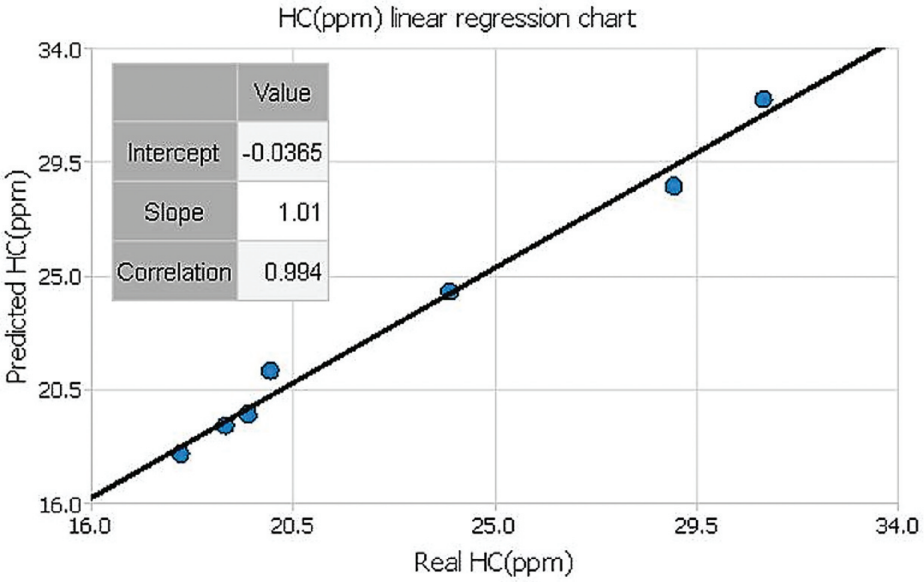


Figure 16. Real HC (ppm) vs. predicted HC (ppm).

rest 20% of data is used for testing instances. Linear regression is performed Using experimental data, and the relation between experimental output parameters and predicted output parameters is established.

Experimental and predicted data validation using ANN

Linear regression analysis is done between the scaled neural network outputs and the associated targets to assess the loss of a model. For each of the output variables, this analysis consists primarily of three factors. The two-parameter connections to the y-intercept and slope of the best linear regression related to objectives and scaled outputs are referred to as a and b. The correlation coefficient, or R^2, is the third parameter. It establishes a link between objectives and scaled outcomes. If the slope is one, we have a perfect fit, meaning the outputs and targets are identical, and the y-intercept is zero. If the correlation coefficient (R^2) is equal to one, the Neural Network results and the targets in the testing subset are perfectly correlated.

Relation between various parameters of the engine with real and predicted values

Estimates based on ANN aiming at the CI engine's performance and emission proved satisfactory statical performance. The graph above illustrates that ANN can simulate a single cylinder CI engine with opo-diesel fuel blends by combining the findings from research and ANN guesstimates. The assessment of ANN guesstimates results and investigational results for BTE in Figure (10) BSEC in Figure (11) and EGT in Figure (12) . According to found data, R^2 values were 0.994, 0.986, and 0.98, respectively. The mean relative error is also calculated for BTE, BSEC, and EGT, and it is 0.864%, 0.568%, and 2.035%, respectively.

The assessment of ANN guesstimates results and investigational results for NOx in Figure (13) Opacity in Figure (14), C.O. in Figure (15) and H.C. in Figure (16) . According to found data, R^2 values were 0.997, 0.993,0.973, and 0.994 respectively. The mean relative error is also calculated for NOx, Opacity, C.O., and H.C., and it is 4.01%, 4.5%, 4.112%, and 3.42%, respectively, and similar trends were seen in various works(Owora 2020; Tran et al. 2021).

Conclusion

In the present experimental study, homogeneous blends of diesel and orange peel oil are prepared, and these blends were tested on a light-duty diesel engine. The physicochemical properties of the blends were comparable with neat diesel. The engine results were validated by performing ANN modeling. The major findings obtained are listed below:

- opo blends have comparable Physico-chemical properties to diesel, such as calorific value, viscosity, and density, and so may be used as a diesel fuel substitute.
- At peak engine load, the BTE of opo30 blend is found to be highest, i.e., 17.5%, as compared to neat diesel. The increase in BTE and subsequent 20% reduction in BSEC is due to improvement in the calorific value with the addition of orange peel oil, and also ANN results are close to each other. R^2 and mean relative error for BTE, BSEC is 0.994 and 0.864%, 0.986 and 0.568% respectively.
- The NOx emission for opo30 is 4.17% higher than diesel due to the improved combustion efficiency and increased in-cylinder temperature. ANN results for EGT and NOx are encouraging. R^2 and mean relative error for EGT, NOx is 0.98 and 2.035%, 0.997 and 4.01%.
- In contrast with diesel, a decrease in emission of 5.8% in HC and a 9.7% decrease in C. O. emissions were observed for opo30, and from ANN modeling, the value of R^2 and mean relative error for H.C., CO is 0.994 and 3.42%, 0.973, and 4.112% respectively.
- The smoke emissions were also reduced by 9.7% compared to neat diesel. Moreover, from the ANN modeling for smoke, the value of R^2 and mean relative error is 0.993 and 4.5%, respectively.

The following recommendations can be made based on the above results:

(1) Orange peel oil, i.e., opo100, should be used neat, and its performance and emissions should be tested over a lengthy period of engine running.
(2) To explore a new window of opportunity, investigation can also be conducted on utilizing the Exhaust gas recirculation technique to regulate the NOx emissions, which were observed to be higher with orange peel oil blends.

Notes on contributors

Sandeep Kumar Barnawal is working as a Research Scholar at the Centre for Advanced Studies and Research in Automotive Engineering, Dept. of Mechanical Engineering, Delhi Technological University, Delhi. He did his B.Tech in Mechanical Engineering from DIT University Dehradun; Pursuing M.Tech in Thermal Engineering from Delhi Technological University. His research interests include; alternative fuels with special emphasis on next-generation biofuels. He is currently working on different fuel reformulation projects such as nano-fuels, recovery & utilization of crude glycerol as fuel for light-duty CI engines.

Naveen Kumar was born in August 1967 in Uttar Pradesh, India; Prof. Naveen Kumar did the B.E. from Dayal Bagh Educational Institute, Agra, the M.Tech from IIT Delhi, and Ph.D. from the University of Delhi. He is a Fellow of Institution of Mechanical Engineers (FIMechE), UK; Fellow of Institution of Engineers (FIE), India and Charted Engineer, Engineering Council, UK. His research interests include; alternative fuels with special emphasis on biofuels, decentralized energy systems, renewable energy, waste recycling, and sustainable development. Presently working as Professor of Mechanical Engineering at Delhi College of Engineering since 1996; Prof. Kumar possesses 29 years of experience in academics, industry, and research. He has published more than 100 research papers in the International Journals of repute and more than 100 in Indian Journals and Conferences.

ORCID

Sandeep Kumar Barnawal http://orcid.org/0000-0003-4161-9566
Naveen Kumar http://orcid.org/0000-0002-4408-9997

References

Brindhadevi, K., Shanmuganathan, R., Pugazhendhi, A., Gunasekar, P., & Manigandan, S. 2021. Biohydrogen production using horizontal and vertical continuous stirred tank reactor- a numerical optimization. *International Journal of Hydrogen Energy* 46 (20):11305–12. doi:10.1016/j.ijhydene.2020.06.155.

"Citrus : World markets and trade." 2021.: 1–13.

Deep, A., et al. 2013. Experimental investigation of orange peel oil methyl ester on single cylinder diesel engine. *SAE Technical Papers* :6. https://doi.org/10.4271/2013-24-0171

Deep, A., Kumar, N., Kumar, M., Singh, A., Gupta, D., & Patel, J. S. 2015. Performance and emission studies of diesel engine fuelled with orange peel oil and n-butanol alcohol blends. *SAE Technical Papers. https://doi.org/10.4271/2015-26-0049*

Dharma, S., Hassan, M. H., Ong, H. C., Sebayang, A. H., Silitonga, A. S., Kusumo, F., & Milano, J. 2017. Experimental study and prediction of the performance and exhaust emissions of mixed jatropha curcas-ceiba pentandra biodiesel blends in diesel engine using artificial neural networks. *Journal of Cleaner Production* 164:618–633. doi:10.1016/j.jclepro.2017.06.065.

Hoekman, S. K., and C. Robbins. 2012. Review of the Effects of Biodiesel on NOx Emissions. *Fuel Processing Technology* 96:237–49. doi:10.1016/j.fuproc.2011.12.036.

Hosseini, S. H., A. Taghizadeh-Alisaraei, B. Ghobadian, and A. Abbaszadeh-Mayvan. 2020. Artificial neural network modeling of performance, emission, and vibration of a ci engine using alumina nano-catalyst added to diesel-biodiesel blends. *Renewable Energy* 149:951–61. doi:10.1016/j.renene.2019.10.080.

Kumar, A. R. M., M. Kannan, and G. Nataraj. 2020. A study on performance, emission and combustion characteristics of diesel engine powered by nano-emulsion of waste orange peel oil biodiesel. *Renewable Energy* 146 (x):1781–95. doi:10.1016/j.renene.2019.06.168.

Manigandan, S., Gunasekar, P., Praveenkumar, T. R., Sabir, J. S. M., Mathimani, T., Pugazhendhi, A., & Brindhadevi, K. 2021. Performance, noise and emission characteristics of di engine using canola and moringa oleifera biodiesel blends using soluble multiwalled carbon nanotubes. *Fuel* 289 (October 2020):119829. doi:10.1016/j.fuel.2020.119829.

Martín, M. A., R. Fernández, M. C. Gutiérrez, and J. A. Siles. 2018. Thermophilic anaerobic digestion of pre-treated orange peel: Modelling of methane production. *Process Safety and Environmental Protection* 117:245–53. doi:10.1016/j.psep.2018.05.008.

Owora, C. U. 2020. Evaluation of emission pattern of compression ignition engines fuelled with blends of orange peel oil based biodiesel using artificial neural network model. *International Journal of Environmental Sciences & Natural Resources* 24 (4):4. doi:10.19080/IJESNR.2020.24.556145.

Purushothaman, K., and G. Nagarajan. 2009. Experimental investigation on a C.I. engine using orange oil and orange oil with DEE. *Fuel* 88 (9):1732–40. doi:10.1016/j.fuel.2009.03.032.

Siva, R., D. B. Munuswamy, and Y. Devarajan. 2019. Emission and performance study emulsified orange peel oil biodiesel in an aspirated research engine. *Petroleum Science* 16 (1):180–86. doi:10.1007/s12182-018-0288-0.

Tomar, M., Jain, A., Pujari, P. C., Dewal, H., & Kumar, N. 2020. Potentials of waste plastic pyrolysis oil as an extender fuel for diesel engine. *Arabian Journal of Geosciences* 13(13). https://doi.org/10.1007/s12517-020-05574-6

Van Hung, T., Alkhamis, H. H., Alrefaei, A. F., Sohret, Y., & Brindhadevi, K. 2021. Prediction of emission characteristics of a diesel engine using experimental and artificial neural networks. *Applied Nanoscience*. doi:10.1007/s13204-021-01781-z.

Uslu, S., and M. B. Celik. 2018. Prediction of engine emissions and performance with artificial neural networks in a single cylinder diesel engine using diethyl ether. *Engineering Science and Technology, an International Journal* 21 (6):1194–201. doi:10.1016/j.jestch.2018.08.017.

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.

Experimental Studies on Frozen Soil

Kshitij Gaur (kshitijgaur_2k21phdce08@dtu.ac.in)
Delhi Technological University https://orcid.org/0000-0003-1424-1430
Anil Kumar Sahu
Delhi Technological University

Research Article

Keywords: Frozen Soil, Soil Experimentation, Frozen Soil Parameters, Frozen Soil Tests

DOI: https://doi.org/10.21203/rs.3.rs-803641/v5

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Various studies have been carried out for soils at normal room temperature but the studies on frozen soils are meagre. For every construction, soil investigation is the most important and the primary step for a site. For constructions at normal room temperature, there is plenty of experimentation and research data on soil is available. But lack of research data for colder regions, where the ambient temperature is below zero degrees Celsius for most of the time. It is therefore the need to study soil under the iced condition to get a better idea about the behaviour of frozen soils. There is little research on the construction and mechanical behaviour of frozen soil but no study on the very basic parameters like void ratio, bulk density, porosity, and the degree of freezing and how these parameters change as the soil temperature changes from normal room temperature to negative values. The main emphasis is on the study and experimentation of frozen soil and the formulation of different relationships between individual soil parameters at various temperatures. The methodology used is to model the soil surface (open grounds in colder regions) by taking sand as the soil after sieving. The model samples are taken into beakers with different bulk densities to replicate real site conditions in the freezer. Then by calculating factors like density, porosity, void ratio, etc at negative temperature (-5, -10, -15, -20 degree Celsius) and forming a relationship with the same parameters as that on room temperature. The experimental data obtained is used in “Eureqa software” that will utilize the input so provided and will find mathematical relations that exist in the soil parameters.

1 Introduction

The world is expanding in terms of development and construction every day. From the construction of high-rise bridges to underwater structures, modern civil engineering is leaving no page unturned. The constructions on glacial soils or frozen soils are of concern as the detailed studies are not available like that of soils under normal room temperature. The glacial soils behave in a different manner than the usual soil at normal temperatures. The properties vary significantly as the temperature changes from normal room temperature to negative values, tested and observed by Baker and Spaans [9]. All the experiments and relations are established for normal soils at room temperature while glacial soils have not been studied in detail as of now. The main objective is to study the behaviour of glacial soils or frozen soil and establish relationships between different characteristics of soil with the variation of temperature, studied by Li et al. [4]. Just for example, when the voids of the soil get increased the chances of inter-connectivity also increase, and thereby permeability of soil increases, observed by Zhu and Carbee [3]. But in the case of glacial soil, the voids get occupied by ice instead of water, which acts as a barrier to the flow of water and thereby permeability decreases. This example gives the idea that the governing factors and the formulae applied to the soil under normal room temperature don’t hold well for frozen soil. The author will obtain actual data of soil under both room temperature as well as under frozen state. Then using those data as an input to “Eureqa” software[18] which will give the hidden mathematical relationship that exists between the parameter under varying temperature. For example, the porosity of the soil sample will be calculated at both room temperature as well as the frozen state. Then

by knowing the water content and temperature of frozen soil the relationship can be established. This will be done by taking frozen soil porosity to be a function of water content, porosity at room temperature, and frozen soil temperature. The Eureqa software will then give the relationship between both the porosity of the same soil sample with temperature variations. This will help the geotechnical engineer to know the different properties of frozen soil if its data is given for room temperature.

The purpose of this research is to study the behaviour of glacial soils or frozen soil and establish relationships between different characteristics of soil, which was observed by Smith et al. [2]. Andersland and Ladanyi [6], performed certain tests and observed that the surface area, expandable clay lattice, and the activity ratio, was deviated from what was assumed. Viklander [5] conducted triaxial drainage shearing on remolded silt samples with different cycle times of freeze-thaw. It was concluded that the cohesion of remolded silt increases compared with that of the non-freeze-thaw group after the freeze-thaw cycle. Janoo and Shoop [1], grouped undisturbed fine-grained soil and remoulded them into saturated remolded soil samples of various degrees, and measured the effects of water content and confining pressure on the mechanical properties of fine-grained remolded soil by applying different confining pressures on them. This article reference in Rock And Soil Engineering analysis method for silty clay, and in strict accordance with the Geotechnical Test Method Standard to test the soil sample. Edwards and Cresser [12], performed testing of the soil sample density, water content, particle composition, the basic parameters such as liquid limit and the plastic limit is important. Through the analysis of Grey Theory, this article finally analyses the shearing strength of the silty clay change rule, provides the reference to the following seasonal frozen area subgrade problems to lay the foundation, observed by Altuhafi and Baudet [13]. Tulaczyk et al. [20], inferred those studies on the behaviour of glacial sediments are limited, and those studies are concentrated only on investigating the ultimate strength of frozen soil and mechanical behaviour. An elaborative model considering the particle breakage and its plastic shear mechanisms is established by Chang et al. [24]. The shear strength and elastic modulus of frozen soil as studied by Khoroshilov et al. [25], observed the shear strength parameters of frozen soil. Also, the influence of freezing and thawing cycles on wind erosion strength of black soil simulations done by Liu et al. [15]. Few other studies also include the three-dimensional compressibility anisotropy of a thawing soil and are determined by the cryogenic constitution of the frozen soil and affect primarily the value of the coefficient of thawing, Bakulin and Zhukov [11]. The mechanical properties of frozen soil, the stress-strains are discussed deeply in Qi et al. [19] and Shen et al. [17]. The books on the chemical and biological properties, Edwards and Cresser [12], Merzlyakov [23] and the creep in frozen soil, Fish [7], Eranti and Lee [14], and Andersland and Ladanyi [16]. Further, various research has been done on flow throw partially frozen soil. For example, hydraulic conductivity was calculated by Azmatch et al. [8] and Burt and Williams [10] in partially frozen soil. Research work is also done on fully frozen soils. For example, the strength and deformation characteristics of fully frozen soil by Xu et al. [22] and the thermal effects and strain rates by Zhu et al. [21]. Previously the research on frozen soil was done for the analysis of the mechanical, physical, biological, and chemical properties, however, the main motive of this study is to obtain the relationship between characteristics of soil like its porosity, void ratio,

and bulk density, at varying temperature, is forming a void in the studies that needs to be fulfilled to get the knowledge about how these characteristics changes as the temperature drops to negative values.

2 Methodology

The soil is mixed with a certain amount of water(i.e., known water content) and then filled in 100ml cylindrical beakers with different bulk densities. Sample represents soil condition in the open ground of colder regions. The sample is placed in a freezer to replicate the temperature conditions of colder regions. In this study, the working temperatures are 20, -5, -10, -15, -20 degrees Celsius which is achieved in a deep freezer in the laboratory. The soil factors like porosity, void ratio, bulk density, dry density, etc are calculated for frozen soil at the above-said temperatures. These are then related to the same properties under room temperature. For instance, the porosity of frozen soil is a function of porosity at room temperature, water content, and the temperature for which it is being calculated. The degree of freezing is also be calculated using the same approach. After getting data of normal and frozen soil, different relations can be found using eureqa software.

The test procedures performed in this experimental work are according to Indian Standard codes: IS 2720-2(1973), (IS-2720-PART-3-1980), ISO 11272:2017. The procedure is as follows:

A) Oven dry the soil at 105 degrees Celsius for 24 hours as per **IS 2720-2(1973)**. Put this soil in plastic bags so that it may not get any moisture from the surrounding.

B) Take out samples from this soil bag. Sieve the soil for classification. For this research, locally available soil is used and sieve analysis is done on that.

C) Let the dry weight of soil sample= W_s

Add known amount of water content (2%,4%,6%,8%,..,40%)

$w= (W_w/W_s)$ ……(1)

D) Separately, determine specific gravity(G_s) of soil solids using a pycnometer test.

And relate it with equation (2) to get the value of γ_s. As per **IS-2720-PART-3-1980.**

$Gs=(\gamma_s/\gamma_w)$ ……(2)

Once, γ_s is calculated volume of soil solids will be further calculated by

$\gamma_s=(W_s/V_s)$ ……(3)

E) Now, calculate the void ratio(e) as per IS : 2720 (Part 2)

$e=(V_v/V_s)$ ……(4)

F) V_v can be calculated by first measuring the total volume(V_t) of soil sample having different bulk densities(γ_{t1}, γ_{t2}, γ_{t3}...) in measuring container as per **ISO 11272:2017** and then using:

$V_v=(V_t - V_s)$ ……(5)

A. Now, place the soil sample container in the freezer for various temperatures to record data (-5, -10,-15,-20 degrees Celsius)

B. After 24 hours, take the sample out and weigh it. Also, measure heights at five different locations, and the average is taken. Do this step as soon as possible to avoid error due to temperature change.

C. Then the ice content of soil sample will be:

$w'=(W_l/W_s)$ ……..(6)

The water content will slowly change to ice content as the temperature is kept on decreasing. Therefore, the ice content is a function of the initial water content and the temperature.

$w'= f(w, T)$ ……..(6A)

This ice content will reach a maximum value when all the water present in the soil gets converted into ice. The void ratio will also change as volume ice expands filling all the voids at first and then increasing the overall volume of the soil sample as shown in **Fig 2.1**.

3 Observations And Results

In this study, the author has studied the following soil parameters and their relationship at various temperatures. The study is performed on locally available soil and the soil gradation curve after sieve analysis suggests that the soil is well-graded sand. The C_u and C_c of the soil found is 8.33 and 1.09 respectively. The gradation curve for the soil is shown in **figure 3.1**.

After performing the experimental studies on frozen soil, the following relationship between parameters of frozen and unfrozen soil at varying temperatures are found.

Weight Function

The weight of frozen soil can be found if, the weight of the same soil at room temperature, its water content and the temperature at which weight is to be determined is known. By entering the data obtained in the laboratory, Eureqa software gives the following relationships (**Fig. 3.2**) The weight function would be:

$W'= f(W,w, T)$ ……..(7)

where w is the water content of soil at room temperature.

W is the weight of soil at room temperature

T is the temperature of frozen soil.

The relationship found is:

$W' = W + 0.992w - 0.000208\ W(-T)$ ……..(7A)

This relationship has a maximum error of 0.18.

Void Ratio

The void ratio of soil will change with temperature as the volume of water gets increased when it changes from liquid state to iced state. Thereby increasing the volume of voids while the volume of soil solids remains the same. Therefore, the void ratio will increase which is confirmed practically too. By entering data obtained in the laboratory, the following relationship is obtained by the eureqa Software (**Fig. 3.3**).

$e' = f(e, T)$ ……..(8)

$e' = e + 0.00503(-T)$ ……..(8A)

Where T is in Degree Celsius.

The graph between percentage change in void ratio vs temperature suggests that at higher water content the change in void ratio is more and a steep curve is obtained as compared to lower water content as shown in figure 3.4. This is because more water will convert into ice and the volume of voids will increase sharply in case of higher water content.

Bulk Density of Soil

The weight of soil solids remains constant but the weight of water changes with the change of temperature thereby changing the bulk density of soil. Therefore, a relationship may exist between the bulk density at room temperature to the bulk density at frozen temperature.

After entering the data obtained in laboratory, the following relationship is obtained by eureqa software(**Fig. 3.5**)

The function is

$\gamma_t' = f(\gamma_t, w, T)$ ……..(9)

$\gamma_t' = \gamma_t - 0.0019(-T)\gamma_t^2$ ……..(9A)

As the graph in Fig. 4.13 suggests that the change in bulk density will be steeper when working with higher water content. Also, the rate of change in bulk density with temperature decreases as the temperature is decreasing. This is because the density of ice is smaller than that of water.

Degree Of Freezing

Analogous to the degree of saturation (S), the degree of freezing (F) would tell the degree up to which the soil is frozen. For example, the soil is fully saturated i.e., all the pores are filled with water. At -2 degrees Celsius, a fraction of water will get crystallized but at -20 degrees Celsius, all the water will get frozen. Therefore, the degree of freezing will give an idea about the degree up to which the water in pores of soil is frozen. The data obtained from the laboratory leads to the following result in eureqa software (**Fig. 3.7**)

$S= (w* G_s)/e$ ……..(10)

The relationship is given by,

$F= f(w,e',T)$ ……..(11)

$F= (2.77\ w/e') - 0.0387(T)$ ……..(11A)

where F is the degree of freezing computed as,

$F= V_l/V_v$ ……..(11B)

Porosity

The porosity of soil is a function of temperature. When the temperature becomes negative, the water present in the pores converts into ice which expands in volume. The increased volume of ice at first tries to occupy all the void space. Once occupied, on further increase in water content, more water turns into ice which expands more than the void space. This thereby increases the volume of voids. When the volume of voids increases, the porosity will also increase. By entering data obtained from the laboratory in the Eureqa software, the following relationship is obtained (**Fig. 3.8**). The function of porosity with temperature can be given by;

$n'= f(n,w,T)$ ……..(12)

$(n') = 0.3376 + 0.0003*T + 0.74*n^2 - 0.00069*w*\cos(234*n^3)$ ……..(12A)

Where,

n'= Porosity of frozen soil

T= Temperature in degree Celsius

n= Porosity at room temperature

The above function has a maximum error of 0.021.

In **figure 3.9**, a graph between the percentage of change in porosity vs temperature is plotted. The trend of the curve suggests that at more water content, the change is more visible marked with a steep slope. This is because more water will convert into ice and the volume of voids will increase sharply in case of higher water content.

After performing the soil parameters analysis on eureqa software above, the author comes across various results as follows:

a) $W' = W + 0.992w - 0.000208\ W(-T)$

Using this formula, one can obtain the water content of frozen soil if the weight of frozen soil, the weight of soil at room temperature, and the temperature of frozen soil are known. This formula thereby helps in calculating the water content of the soil sample theoretically.

b) $e' = e + 0.00503(-T)$

Using this formula, one can obtain the void ratio of frozen soil if the void ratio of the same soil sample at room temperature and the temperature of frozen soil is also known.

c) $\gamma_t' = \gamma_t - 0.0019(-T)\gamma_t^2$

Using this formula, one can calculate the bulk density of frozen soil sample, if the bulk density of same soil sample under room temperature is known along with the temperature at which frozen bulk density is required.

d) $F = (2.77\ w/e') - 0.0387(T)$

Using this formula, one can calculate the degree of freezing which will indicate the amount of freezing that occurred in the frozen sample. This term is analogous to the degree of saturation in normal room temperature soil which tells the amount of saturation of soil sample.

e) $(n') = 0.3376 + 0.0003*T + 0.74*n^2 - 0.00069*w*\cos(234*n^3)$

Using this formula, one can obtain the porosity of frozen soil if the porosity of same soil sample at room temperature is known and the temperature of frozen soil is also known.

Where,

W'= Weight of Frozen soil.

W = Weight of soil at room temperature.

T = Temperature at which soil is being investigated.

W = Water content of the soil.

γ_t' = Bulk unit weight of frozen soil.

γ_t = Bulk unit weight of soil at room temperature.

F = Degree of freezing.

n = Porosity of soil at room temperature.

n' = Porosity of frozen soil.

4 Discussion

There are five relationships that have been established between unfrozen and frozen soils. Formulae so obtained are valid up to -20 degrees Celsius and are tested on sandy soil. Below this temperature, the soil parameter becomes constant. Therefore at -20 degrees Celsius, the soil becomes fully frozen, and the water present in the voids of the soil is fully transformed into ice by going through the crystal phase. The leads of this study can help other authors for detailed study of frozen soils and to take the results of this research to dig further into the newly studied branch of frozen soil mechanics.

The Novelty of this paper lies in the methodology of the work performed, as this includes a new idea of relating the properties of frozen soil with that of unfrozen soil. From these relationships, one can found properties of frozen soil if the data of unfrozen soil is available and vice-versa. While all the research that has been done on frozen soil prior to this is about mechanical, physical, biological, and chemical properties but the basic idea about the relationship between characteristics of soil like its porosity, void ratio, bulk density, etc at different temperatures is forming a gap in the studies that need to be fulfilled to get the knowledge about how these characteristics change as the temperature drops to negative values.

In various regions of Northern India, the temperature goes as high as 20 degrees Celsius but for a small period, while throughout the major time it remains negative. For such sites and other colder regions, if data is obtained for soil at higher temperature (say, 20 degrees Celsius), then when the same soil becomes frozen then the relationships developed in this research would be helpful to determine the characteristics at frozen temperature.

For example, if the value of the void ratio is 0.6 at 20 degrees Celsius, its water content is 25% and it is required to calculate the void ratio and degree of freezing of the same soil at the same density at -20 degrees Celsius. To calculate this, one needs not to measure the void ratio of frozen soil if he have a handy formula with him.

e’= e + 0.00503(-T) ……..(from 8A)

And,

F= (2.77 w/e’) - 0.0387(T) ……..(from 11A)

Therefore, the value of void ratio at -20 degrees Celsius of the given soil sample will be 0.7006 and the degree of freezing will be 99.617% which means at -20 degree Celsius, almost all the water content is converted into ice and the soil tends to be fully frozen.

Likewise, formulae for porosity, the weight of soil, bulk density developed in this research will help the site engineers to have reliable formulae to measure these parameters of frozen soil if unfrozen soil data is available. The formulae will also help in determining several other functions that use these parameters.

5 Conclusion

- The formulae obtained will help to cover the gaps in the literature review by providing basic information about the frozen soil and its parameters.
- This study can be used for the practical purpose as the data obtained by the author in the laboratory is true and self-verified. The constructions done in colder regions can use the formulae so obtained to carry out soil investigation and understand the frozen soils in a better way. The results so obtained will help to know any parameter of soil at different freezing temperatures if their values at normal room temperature are known or vice-versa.
- This study can play a vital role in determining the parameters of frozen soil and can also prove helpful in further studies of frozen soils.
- This research inculcates all the references and thorough reviewing of the literature reviews related to the field of frozen soil.

6 Declaration

Acknowledgments. The authors would like to thank all the teachers and laboratory assistants in Delhi Technological University for helping and providing all necessary types of equipment required in the successful completion of this study.

Conflict of Interest. On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Janoo, V.; Shoop, S. Study of shear strength of silty clay under freeze-thaw cycle. Earth Environment Science. **2019.** 330 042009.
2. Smith; Chen; Haofeng. Lower and Upper Bound Shakedown Analysis of Structures With Temperature-Dependent Yield Stress. Journal of Pressure Vessel Technology.**2010.** 132. 10.1115/1.4000369.
3. Zhu, Y.L., Carbee, D.L. Creep behavior of frozen silt under constant uni-axial stress In: Proc., 4th Int. Conf. on Permafrost, Fairbanks, Alaska, pp. **1984**. 1507–1512.
4. Li, D.W., Wang, R.H., Zhao, Y.H., Hu, P. Research on parabolic yield-surface creep constitutive model of artificial frozen soil. Rock Soil Mech. **2007**. 28(9), 1943–1948.
5. Viklander, P. Influence of cycles of freezing and thawing on the permeability in soils. Lulea University of Technology.**1995**. SE-971 87.
6. Andersland, O.B., Ladanyi, B. An introduction to frozen ground engineering. Chapman and Hall, New York. **1984**.
7. Fish, A.M.: Strength of frozen soil under a combined stress state. In: Proc., 6th Int. Symp. on Ground Freezing, Beijing, China, **1991**. vol. 1, pp. 135–145.
8. Azmatch, T.F., Sego, D.C., Arenson, L.U., and Biggar, K.W. Using soil freezing characteristic curve to estimate the hydraulic conductivity function of partially frozen soils. Cold Reg. Sci. Technol. **2012**. 83–84: 103– 109.
9. Baker, J.M., and Spaans, E.J.A. Mechanics of meltwater above and within frozen soil. In: CRREL Spec. Rep. 79. US Army Corps of Eng., Cold Regions Res. Eng. Lab., Hanover, NH.**1997.** p. 31– 36.
10. Burt, T.P., and Williams, P.J. Hydraulic conductivity in frozen soils. Earth Surf. Processes.**1976.** 9: 411– 416.
11. Bakulin, F.G. and Zhukov; Deformations of Frozen Dispersed Soils during Thawing; **1955**.
12. Edwards, A. C., Cresser, M. S.; Freezing and Its effect on Chemical and Biological Properties of Soil; **1992**.
13. Altuhafı, F.; Baudet, B.. A hypothesis on the relative roles of crushing and abrasion in the mechanical genesis of a glacial sediment. Engineering Geology. **2011**. 10.1016/j.enggeo.2011.03.002.
14. Eranti, E., and Lee, G. C.. Cold Region Structural Engineering. USA: McGraw Hill.**2000.**
15. Liu, Tiejun; Xu, Xiangtian; Yang, Jie. Experimental study on the effect of freezing-thawing cycles on wind erosion of black soil in Northeast China. Cold Regions Science and Technology. **2017**. 10.1016/j.coldregions.
16. Andersland, O. B., and Ladanyi, B.; Frozen Ground Engineering. New York: Wiley; **2004**.
17. Shen, Mingde; Zhou, Zhiwei; Zhang, Shujuan. Effect of stress path on mechanical behaviors of frozen subgrade soil. Road Materials and Pavement Design. **2021**. 10.1080/14680629.2020.1869583.
18. Eureqa Software Created by: Michael Schmidt and Hod Lipson. **2009.**

19. Qi, Jilin; State-of-art of research on mechanical properties of frozen soils; Springer Series in Geomechanics and Geoengineering, pp. 129-133, **2010.**
20. Tulaczyk, Slawek; Kamb, W.; Engelhardt, Hermann. Basal mechanics of Ice Stream B, West Antarctica 1. Till mechanics. Journal of Geophysical Research. **2000**. 10.1029/1999JB900329
21. Zhu, Zhi-Wu; Ma, Yue; Zhang, Hai-Dong; Song, Weidong; Gan, Yuan-Chao. Evaluation of thermal effects and strain-rate sensitivity in frozen soil. Thermal Science. **2014**. 10.2298/TSCI1405631Z.
22. Xu, Xiangtian; Lai, Yuanming; Dong, Yuanhong; Qi, Jilin. Laboratory investigation on strength and deformation characteristics of ice-saturated frozen sandy soil. Cold Regions Science and Technology. **2011**. 10.1016/j.coldregions.2011.07.005.
23. Merzlyakov, V. Physical and Mechanical Conditions for Primary Frost Crack Formation. Soil Mechanics and Foundation Engineering. **2016**. 10.1007/s11204-016-9389-1.
24. Chang, Dan; Lai, Yuanming; Gao, Jianqiang. An investigation on the constitutive response of frozen saline coarse sandy soil based on particle breakage and plastic shear mechanisms.Cold Regions Science and Technology. **2018**. 10.1016/j.coldregions.
25. Khoroshilov, K., Katarov, V.; Gavrilov, Timmo; Kolesnikov, Gennady. Strength and Elastic Modulus of Frozen Sandy Soil as a Material for a Logging Road. Resources and Technology. **2019**. 10.15393/j2.art.2019.4742.

Figures

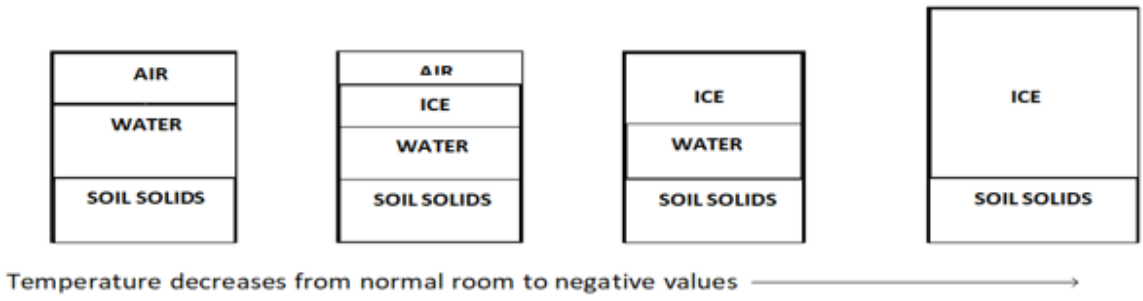


Figure 1

4-Phase diagram of soil

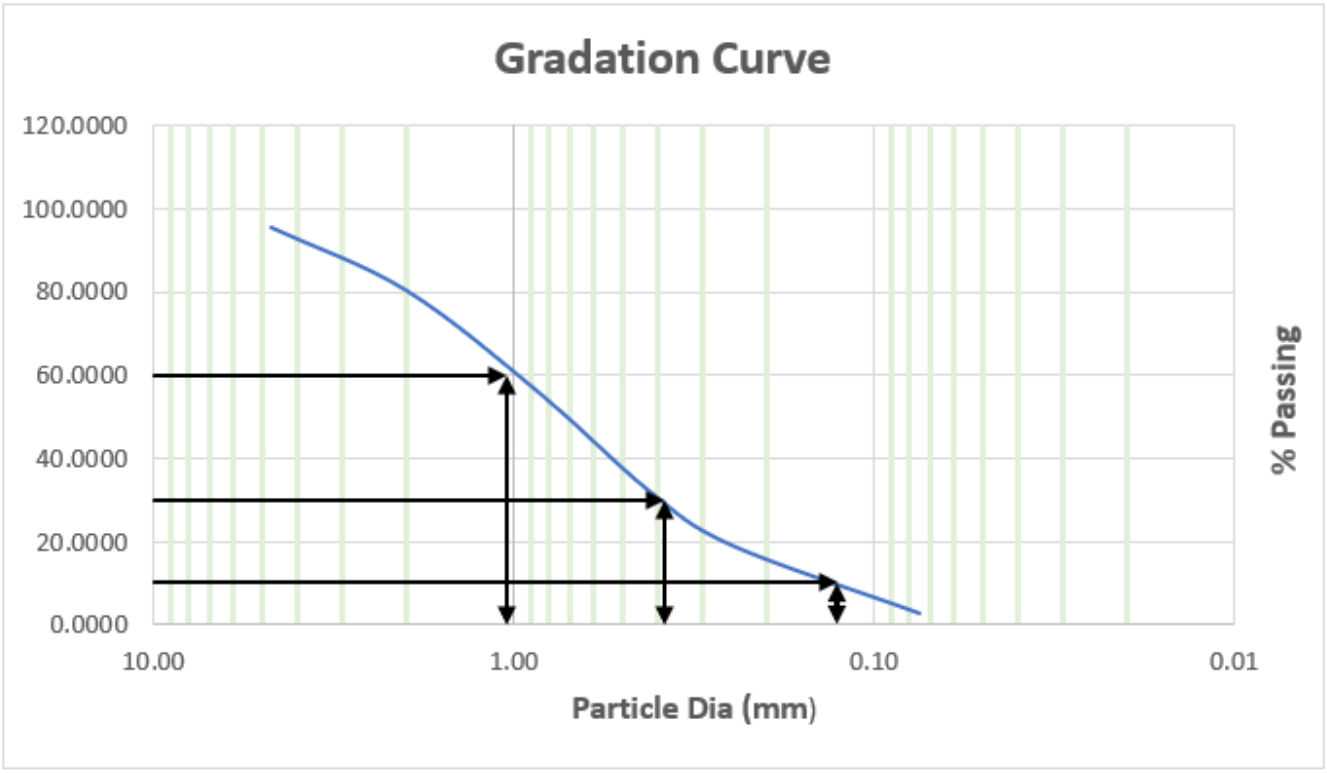


Figure 2

Gradation curve for the soil.

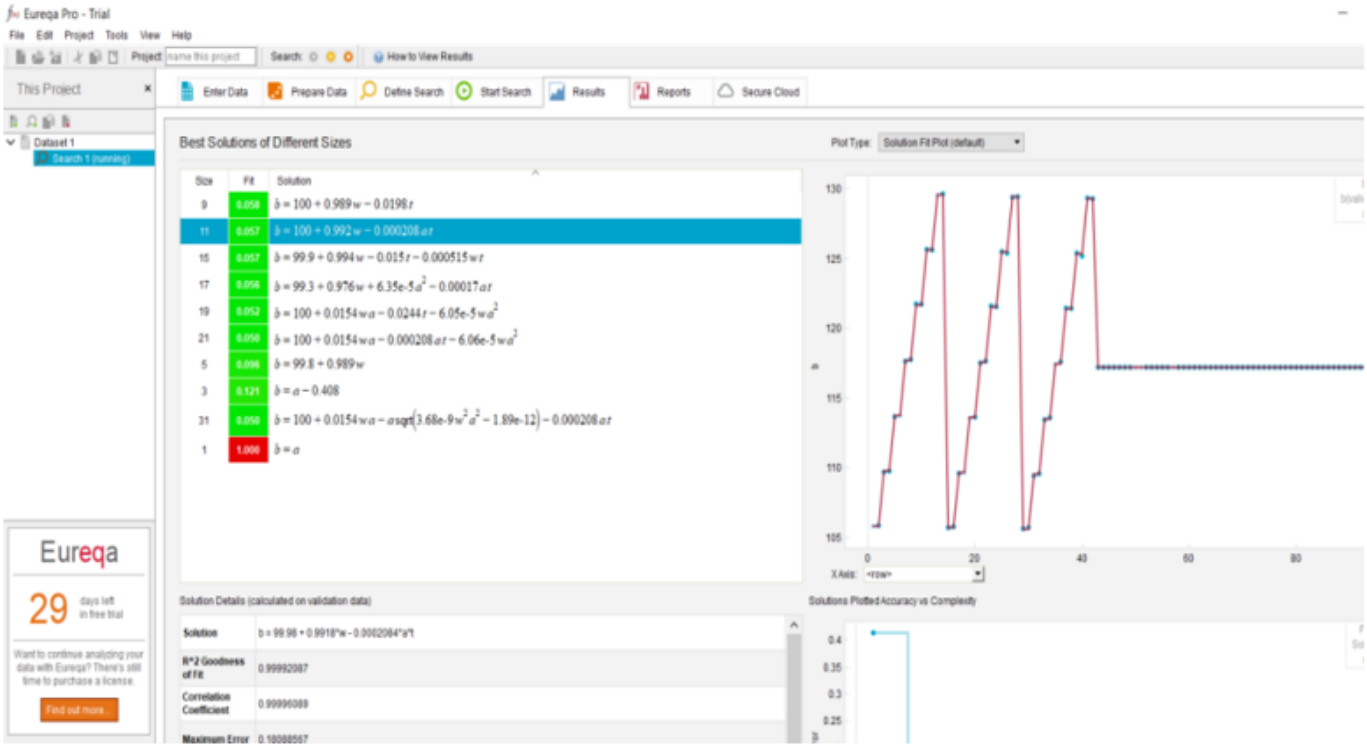


Figure 3

Weight relationship

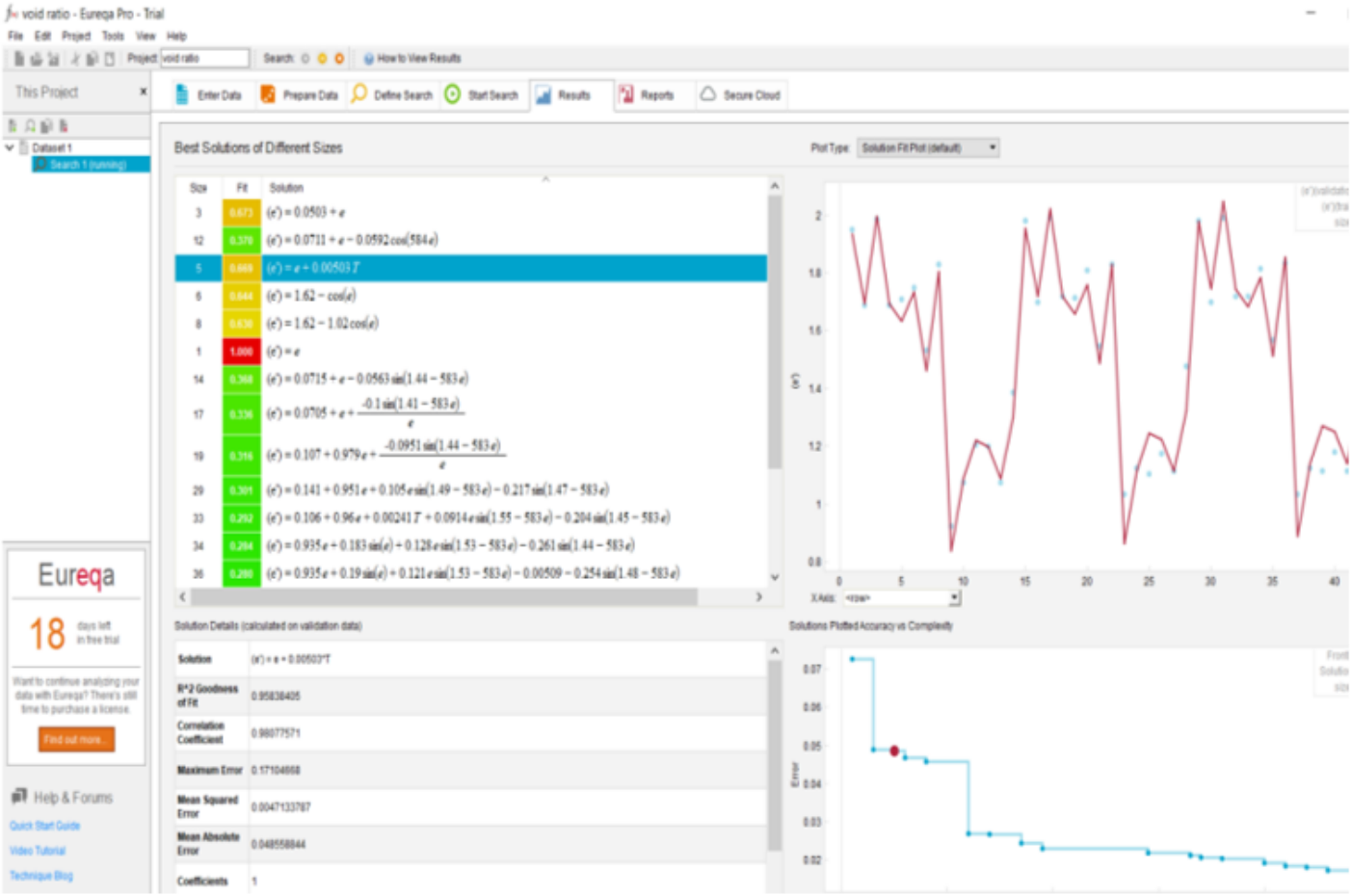


Figure 4

Void ratio relationship.

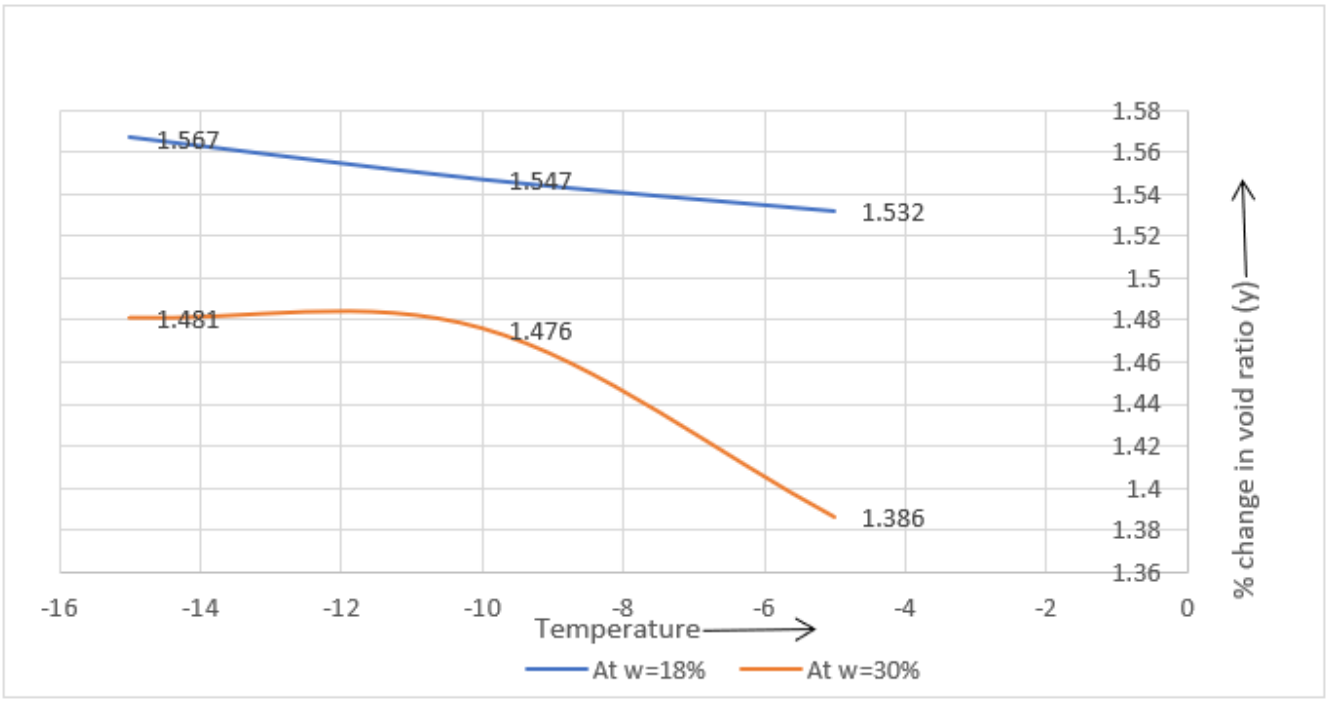


Figure 5

Graph of % the change in void ratio (y) vs Temperature

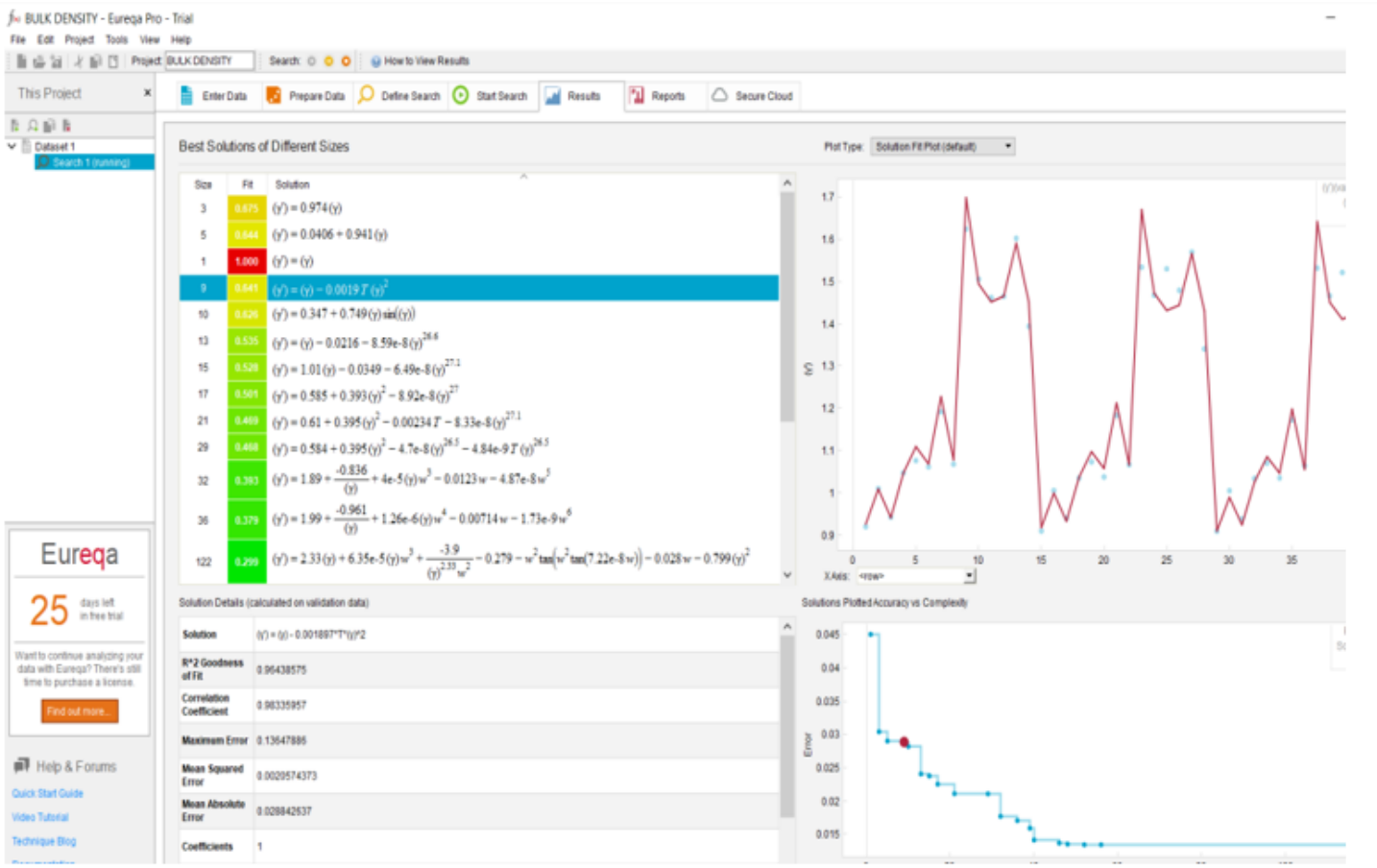


Figure 6

Bulk density relationship.

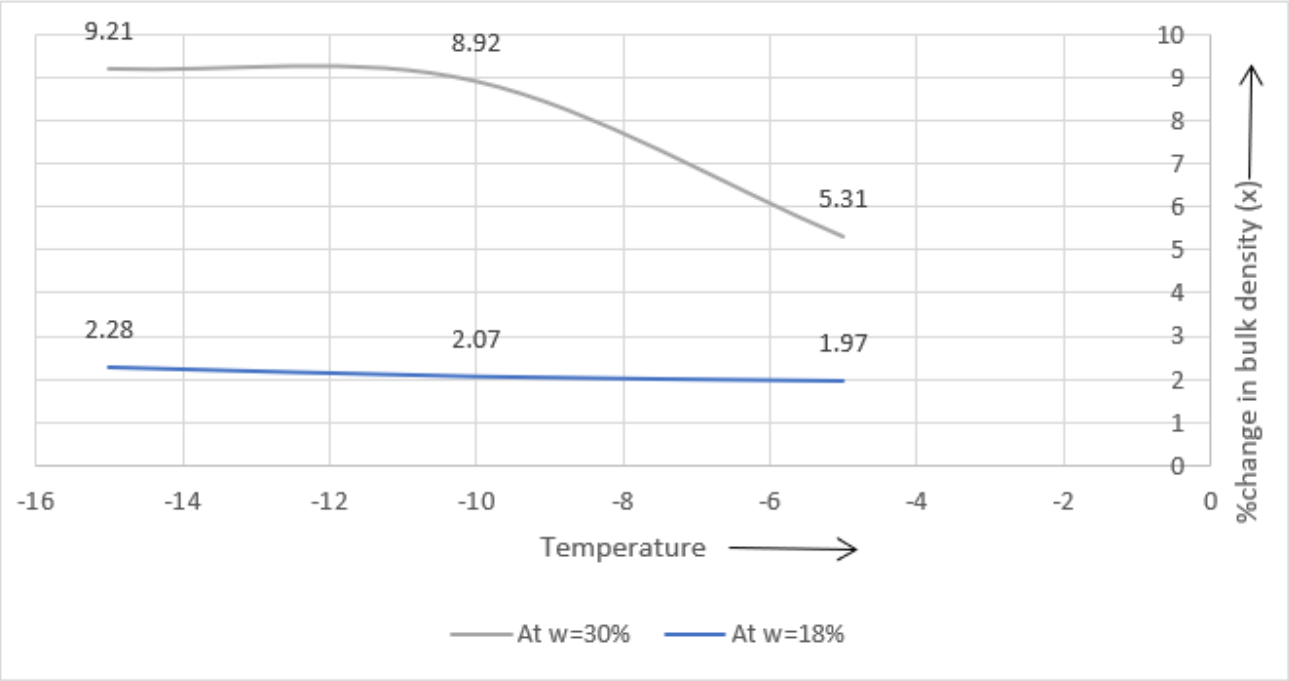


Figure 7

Percentage change in Bulk density Vs Temperature

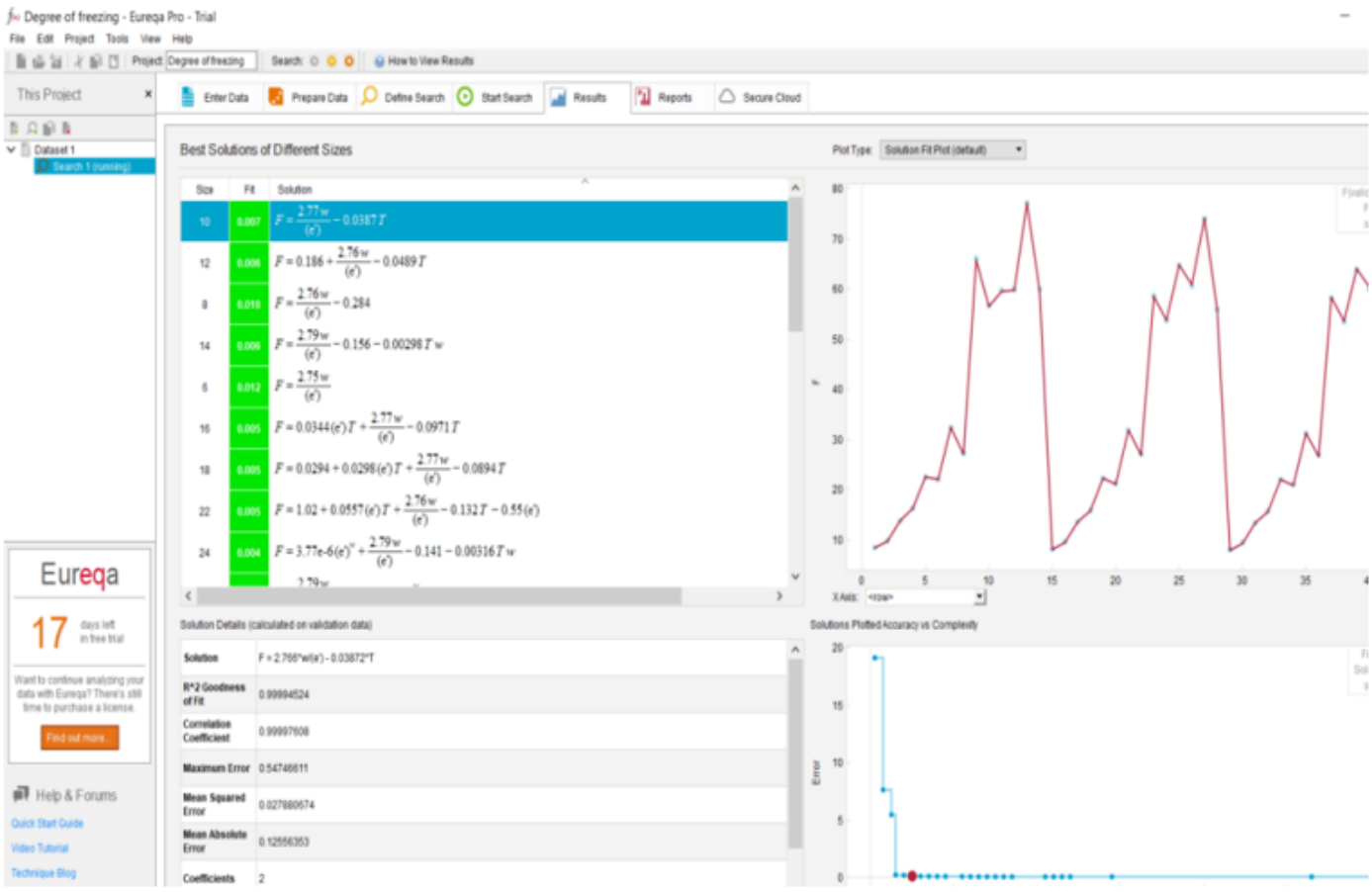


Figure 8

Degree of freezing relationship

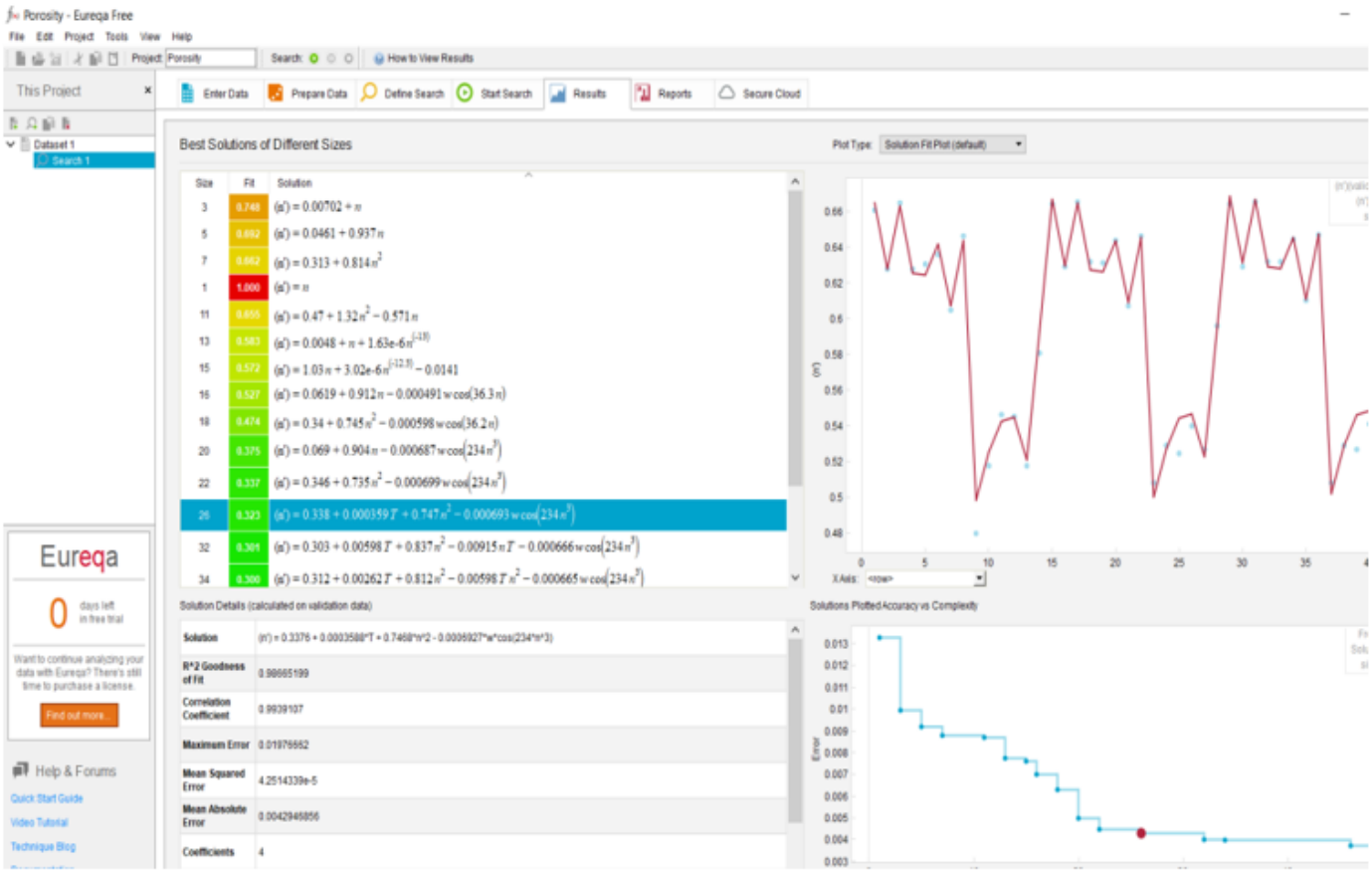


Figure 9

Porosity relationship.

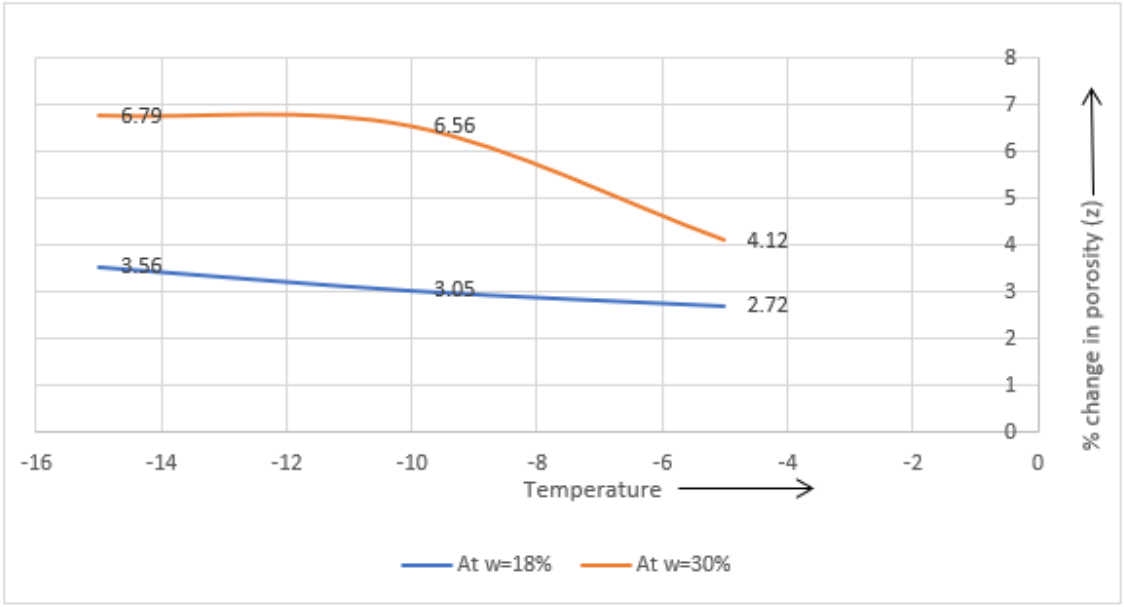


Figure 10

Graph of % change in Porosity (z) vs Temperature.

International Journal of Precision Engineering and Manufacturing
https://doi.org/10.1007/s12541-021-00565-3

Online ISSN 2005-4602
Print ISSN 2234-7593

REGULAR PAPER

Experimental, Computational and Mathematical Analysis of Hybrid Abrasive Flow Machining Process

Sachin Dhull[1] · R. S. Walia[1] · Qasim Murtaza[1] · M. S. Niranjan[1]

Received: 5 November 2020 / Revised: 18 June 2021 / Accepted: 5 July 2021
© Korean Society for Precision Engineering 2021

Abstract

The current scenario of industrialization requires need for higher productivity which is met by advanced material removal process, i.e., abrasive flow machining (AFM) in which the internal surfaces of the workpiece is machined to higher accuracy level with the help of abrasive laden media. In this paper, the conventional AFM setup has been made hybrid using electrolytic and magnetic force arrangement alongwith rotational effect in order to achieve better results in terms of material removal and surface roughness. The newly developed in-house polymer media were utilized in the process and the input parameters taken during experimentation were magnetic flux, electrolytic rod size and shape, rotational speed, polymer media, abrasive particles and extrusion pressure. It was found that the material removal and surface roughness improvement were more in electrochemo magneto rotational AFM process compared to conventional AFM process. The experimental values were in confirmation with those obtained in the optimization techniques applied, i.e., Taguchi L9 OA, Matlab fuzzy logic and GRA-PCA. In addition, the hybrid mathematical model was developed and effect of different forces occurring in the process and computational flow analysis of media have been explained.

Keywords ECMR-AFM · Hybridization · Material removal · Modelling · Polymer media

Abbreviations

α	Angle between magnetic field and plane of solenoid coil
B_p	Peak magnetic field
ω	Angular frequency of current
M	Mutual induction
R	Erosion rate of workpiece material
V_A	Volume removed by single particle
F_a	Axial force
F_r	Radial force
F_t	Tangential force
m	Mass of abrasive particle
ρ	Density of abrasive material
r	Radius of iron oxide particle
p	Extrusion pressure
p_m	Magnetic pressure
τ_s	Shear stress
V_*	Shear velocity
u_{max}	Maximum fluid velocity
$\frac{du}{dy}$	Velocity gradient of fluid
$\frac{H^2}{8\pi}$	Magnetic pressure
B	Magnetic field
Ø	Magnetic flux
N	Number of abrasive particles
h_i	Response of node of hidden layer
w_i	Weight of connection to hidden layer
f	Transfer function
AFM	Abrasive flow machining
EDM	Electric discharge machining
ECM	Electrochemical machining
MAF	Magnetic abrasive finishing
AECG	Abrasive electro chemical grinding
UAAFM	Ultrasonic assisted abrasive flow machining
MRP	Magneto rheological polishing
ECP/ECF	Electrochemical polishing/finishing
DOA	Dio-iso octyl adiabate

✉ Sachin Dhull
sachindhull1989@gmail.com

R. S. Walia
waliaravinder@yahoo.com

Qasim Murtaza
qasimmurtaza@dce.ac.in

M. S. Niranjan
mahendraiitr2002@gmail.com

[1] Delhi Technological University, Delhi, India

Published online: 23 August 2021

DOP	Dio-iso octyl pthalate
ECMR-AFM	Electro chemical magnetic rotational AFM
RSM	Response surface methodology
ANOVA	Analysis of variance
GRA	Grey relational analysis
GRC	Grey relational coefficient
MPCI	Multi performance characteristics index
FIS	Fuzzy inference system
OA	Orthogonal array
GA	Genetic algorithm
MR	Material removal
SBR	Styrene butadiene rubber
NR	Natural rubber
NTR	Nitrile rubber
PBS	Polyborosiloxane
SLM	Selective laser melting
SR	Silicon rubber
MS	Mild steel

1 Introduction

Abrasive flow machining is the process of fine finishing of internal surface of workpiece using ploughing action of abrasive particles onto the work surface. The polymer media is used to carry abrasive particles embedded in it and it flows past the workpiece under high extrusion pressure provided by the moving piston inside the cylinder filled with media. According to the direction of motion, abrasive flow machining (AFM) process can either be one-way, two-way or orbital AFM. If a machine set up comprises only single set of machine cylinder and media cylinder and the media is pressurized only in single direction it is known as one-way process, whereas in two-way process, the movement of media occurs in two directions due to the pressure differences, while small vibrational force is given to the job in case of orbital AFM process. The abrasive flow machining process is used for machining titanium alloys, ceramics, etc. and its application scope is wide as it is used in biomedical, aviation industries and also in manufacturing of automotive parts like gears, piston rings. In this paper [1], authors have done extensive parameters analysis, developed media and also modeled the process to check the efficiency in terms of material removal. The optimization of output responses was done by the author using artificial viral intelligence method [2]. The authors have analysed the temperature effects on the final results in AFM process alongwith comparison of results used with abrasive particles like silicon carbide, alumina and boron oxide [3]. The influence of the machining variables on GFRP composites was investigated [4]. The parameters were cutting speed and flow rate to optimize material removal, roughness, roundness and cylindricity of workpiece results. The application of abrasive jet machining and hard turning processes have been well explained by authors [5, 6] and AJM process of Al6063/B_4C/$ZrSiO_4$ composite materials is done [10]. The effect of electric discharge machining process on material removal on different types of materials and different optimization results have been elaborated [7–9].

The surface integrity on outer, inner and side surface of workpiece was analysed in terms of machine parameters and finishing mechanism was statistically analysed alongwith metallurgically examination of surface integrity [11]. The conformal channels, i.e., straight and helical were fabricated by Selective Laser Melting (SLM) and internal surfaces were finished by AFM and surface integrity was obtained by variation microscopy technique [12]. The repositioning technique was employed in AFM process to calculate change in height which resulted in determination of gauge material removal by taking difference of corresponding values before and after experimentation [13]. A model was developed to predict material removal distribution along flow direction by meshing surface profile before and after AFM process [14]. The surface roughness was reduced from 120.12 nm to 6.48 nm with mirror surface using shear-thickening newly developed guar gum hyrogel based polymer media [15]. The surface finish (Ra) percentage improvement of 72.7% was obtained by using magnetic abrasives in magneto abrasive flow finishing of aluminium tubes [16]. The Ra value was improved by 87.9% in newly developed abrasive rotary flow polishing method of K9 optical glass [17]. The analytical model to correlate Ra with material removal was developed which resulted in 75.7% decrease in surface roughness in Magnetic Abrasive Finishing (MAF) process [18]. A detailed review of the hybrid processes of AFM is done, i.e., rotational, magnetic, electrochemical, helical, spiral and centrifugal force based AFM, etc. and the output results were explained in detail [19]. The conventional AFM machine setup was made variant in terms of introduction of magnetic forces in the present setup and incorporating the effect of centrifugal forces, so as to get more efficient process in terms of material removal [20].

The different types of hybrid AFM processes are explained in the section in detail. The hybridization of conventional process is done by combining it with either electrochemical, electric discharge or rotational process in order to improve the machining performance. The helical AFM process includes drill bit in the inner side of workpiece and the drill is held stationary. When the polymer media flows past the surface, the helical shape of drill bit gives the flow direction to the media, ensuring proper intermixing and attack of abrasive particles onto the surface. The helical and spiral flow of media occurred due to the incorporation of designed passageway and drill-bit inside the workpiece [21, 22]. The ultrasonic assisted abrasive flow machining (UAAFM) is a new hybrid process in which ultrasonic

vibration is given to the workpiece in perpendicular to the media flow direction [23]. This resulted in more material removal (MR) and better integrity of work surface. In magnetic assisted AFM process, magnetic field is applied around the full length of cylindrical workpiece [24]. In electrochemical abrasive flow machining process, the abrasive media is mixed in electrolyte solution and extruded through the inner surface of workpiece assisted by electrolytic force. The electrochemical discharge machining (ECDM) process was used to micro-machine the glass material and its other applications include micro reactors, micro fuel cells, etc., in which reduced overcut of hole and aspect ratio was improved [25]. The taper angle of profile generated was lesser if tool taper angle was reduced in electrochemical micro machining process (ECMM) [26]. In abrasive assisted electrochemical process, surfaces obtained were smoother as compared to that obtained by laser and electric discharge machining that produced heat affected [27]. As compared to conventional Cu electrodes, Cu-ZrB_2 electrodes showed more MRR, less tool wear due to high melting point differences [28]. In the study of MR fluid flow behavior, at different magnetic flux densities, the rheology graphs were plotted to analyse the flow pattern of the bidisperse fluid [29]. It showed that bidisperse fluid has more yield strength and viscosity than monodisperse fluid. If Magneto Rheological Polishing (MRP) fluid was conditioned after several cycles of finishing operation, then it was forced to flow, otherwise the already stiffened ball end of fluid continuously flow towards the tool tip [30]. Hence it can be concluded that the material removal and surface roughness improvement can be done in many ways, i.e., by use of magnetic, rotational and electrolytic forces. The use of magnetorheological fluid and different types of polymer media also a played a significant role in controlling the amount of material removal in different variants of AFM processes.

On thorough scrutiny of the published research work on magnetic field assisted abrasive flow machining (AFM) and electrolytic finishing processes, it has been observed that the finishing forces during processing has not been controlled by external means. Hence, material removal as well as surface roughness has not been controlled during processing. So, sufficient efforts are required towards the study of combined electrochemical and magnetic field assisted abrasive flow machining (AFM). Several mathematical and empirical models have been developed pertaining to prediction of material removal and surface finish, however most of them are specific and can't be generalized. Hence, comprehensive modeling and simulation of Electro Chemical Machining (ECM) process is necessary. The enhancement of performance of ECM in terms of surface finish and material removal using magnetic field have been demonstrated but many of them need complete exploration. In the present paper, the authors have fabricated hybrid electrochemo magneto rotational abrasive flow machining setup and experimentation was done using newly developed polymer media based on Styrene Butadiene Rubber (SBR), Natural Rubber (NR), Nitrile Rubber (NTR), Polyborosiloxane (PBS) and Silicone Rubber (SR). Response Surface Methodology (RSM) technology was employed to arrive at the result that which of the developed media suits best in terms of material removal. Then experimentation design was developed based on Taguchi L9 Orthogonal Array (OA) and experiments were performed on the developed hybrid machine setup. Then the output results, i.e., MR and Ra were analysed in detail, i.e., effect of magnetic flux, ECM rod size and voltage on output results were analysed graphically. Further, Matlab Fuzzy Logic Optimization, Gray Relational Analysis (GRA) approach were also utilized and the results were compared so as to arrive at the best optimized result. Apart from experimentation, the mathematical and computational modeling of the developed Electro Chemical Magnetic Rotational AFM (ECMR-AFM) setup is also done in detail. The voltage, magnetic flux, flow parameters and cutting forces are calculated and computational flow dynamics of developed ECMR-AFM fixture is also done, alongwith various physical models and algorithms.

2 Experimental Analysis

The authors have done extensive experimental analysis so as to arrive at conclusive results regarding material removal and surface finish improvement of the workpiece on the newly designed and fabricated hybrid ECMR-AFM setup. Basically, in abrasive flow machining set-up, there are two sets of piston cylinder arrangement workpiece fixture and polymer media laden with abrasive particles. The material removal occurs because of the to and fro motion of the media through the hollow workpiece fixed in the machine fixture.

2.1 Experimental Setup Description

The section gives a detailed explanation of ECMR-AFM setup used in the experimentation process. The traditional AFM set up is made hybrid in order to achieve more effectiveness of the process. Firstly, the review of hybrid machine setup in terms of electrolytic, rotational and magnetic forces is done in brief and then the machine setup developed and fabricated in the present study is explained in details. In ECM process using mineral water electrolyte, green machining, low cost, high efficiency electrochemical process was explained, polarization cure was measured in potentiostat of 300 V, working electrode (WE) was workpiece i.e. SUS304, reference electrode (RE) is Ag/AgCl, counter electrode (CE) is tool i.e. WC rod of diameter 0.3 mm, as shown in Fig. 1a. It was concluded that short pulsed supply was rapid

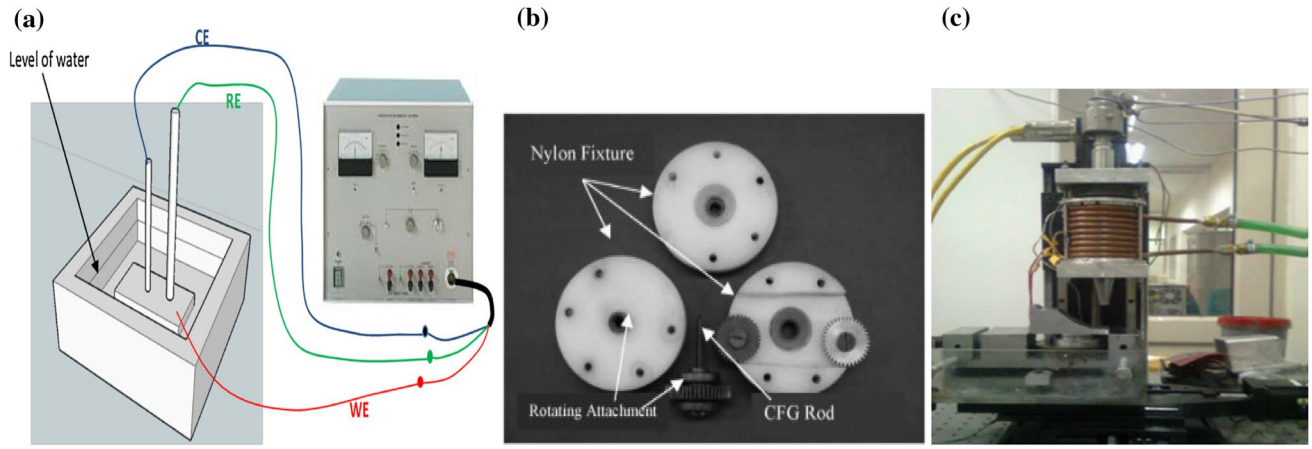


Fig. 1 **a** Electrochemical cell and polarization system [22], **b** CFG rod and nylon fixture arrangement [33], **c** MR finishing setup [29]

than electrolyte $NaNO_3$ [31]. In electrochemical discharge machining (ECDM), aluminium frame was used for machine tool, worm and worm gear speed reduction of 25:1 was obtained, worm was attached to stepper motor shaft [32].

The rotational or centrifugal force assisted AFM process fixture is as shown in Fig. 1b. The machine set up comprised digital oscilloscope, current probe, passive voltage probe were used for monitoring discharge voltage and current waveforms in case of magnetorheological AFM process, as shown in Fig. 1c. In order to obtain fine finishing surface on flat workpieces, electrochemical actions effect was studied by Dabrowski et al. [34]. Centrifugal-magnetic force assisted AFM was developed by Singh et al. [35].

In rotational setup, as shown in Fig. 2, different speeds of rotation of workpiece fixed inside long hollow nylon fixture attached inside the motor with the help of taper bearings were controlled using frequency meter. The advantage of rotation is that it ensures proper mixing of abrasives and

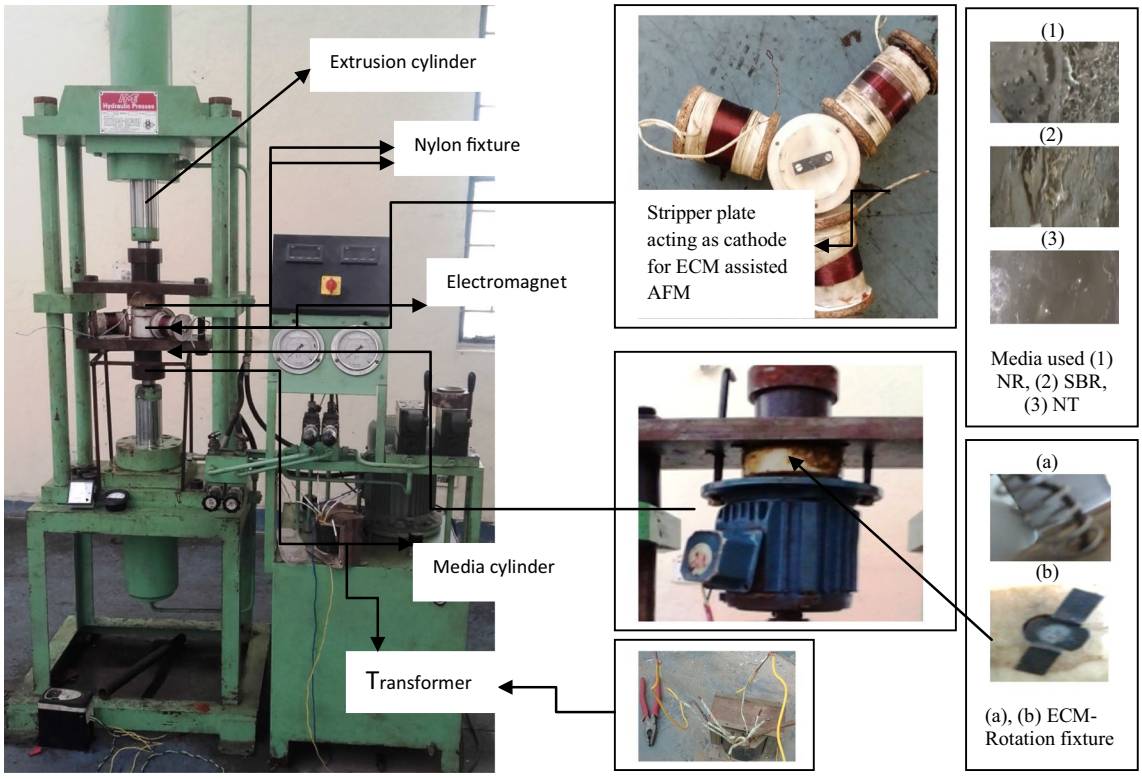


Fig. 2 Hybrid ECMR-AFM set up

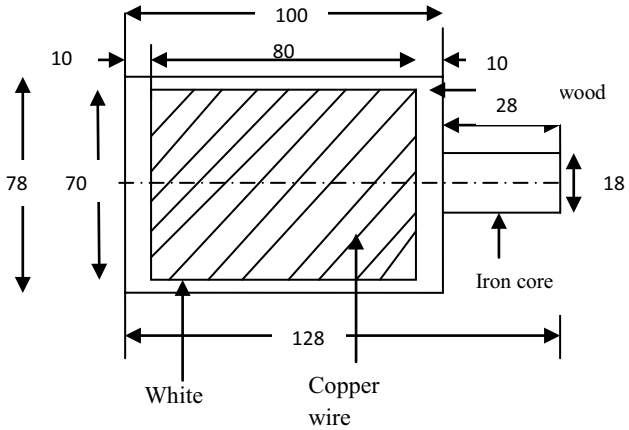
magnetic particles. In order to achieve more effectiveness of the process, an electrochemical force aid has been provided by the authors in the current AFM set up. It has been fabricated in house and a separate power supply transformer was prepared according to the process need. The supply voltages were 6 V, 12 V and 18 V respectively and the electrolyte taken was normal kitchen salt. The workpiece acts as anode since more heat, approximately 67% heat is generated resulting in material removal. The ECM rod of different shapes viz. square, round, helical was placed inside hollow workpiece. The rod acted as cathode in the set-up. Apart from the electrolytic action, the effect of magnetic pull was also analysed by incorporating the electromagnetic solenoid setup arranged along the periphery of job inside the fixture. As shown in Fig. 2, the electromagnets are arranged around the fixture which create magnetic pull on the magnetic particles mixed in the media. The detailed dimension of the electromagnetic solenoid is as shown in Fig. 3a.

2.2 Polymer Media and Abrasive Particles Used in Ecmr-Afm Process

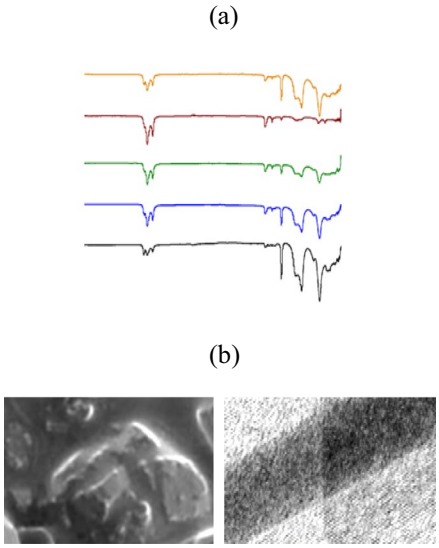
The base polymer i.e., natural rubber (NR), styrene butadiene rubber (SBR), nitrile rubber (NR), polyborosiloxane (PBS) and silicone rubber (SR), alongwith plasticizers used were employed in experimental work Firstly these media were prepared in the laboratory and the plasticizer was mixed in polymer in two roll mill with friction ratio 10:25. The plasticizers used were paraffin oil, dio-iso octyl pthalate (DOP) and dio-iso octyl adiabate (DOA), that resulted in softening the polymeric media and make it flowable since the flowability property would be used in machining process of the workpiece. The preparation of all the polymers was done at two roll mill machine. For preparing the flexible polymers mass, a fixed quantity of polymer was taken first and an appropriate quantity of plasticizer was added.

When the natural rubber sheet passed through the very narrow passage, it was deformed severely and the process kept on going until it became little bit soft. Ultimately, when all plasticizer got consumed, we got a non-sticky semi solid mass, which was ready to get mixed with the gel. The prepared polymer media used is as shown in Fig. 2. When the prepared sample of polymer media is installed on the Fourier Transformation InfraRed/ Thermal Gravimetric Analysis (FTIR/TGA) apparatus, the infrared light falls on sample and the spectrum i.e. transmittance shows fingerprint which gives performance and quality of polymer sample media. The percentage transmittance varied from 0 to 100, whereas wavenumber decreased from 4000 to 500. The FTIR image or graph can be understood by dividing the whole region of the graph into different zones. The media characterization and microsocopic images are shown in Fig. 3b. The different combinations of the variants of AFM processes, i.e. magnetic, electrolytic and rotational force assisted AFM process are depicted in Fig. 4.

In Fig. 5, a block diagram of the prepared machine setup is explained in detail comprising the interface of magnetic field and electrochemical setup with the basic piston cylinder AFM setup. This diagram is a simple way to understand the actual practical machine setup used in the experimentation,



(a) Electromagnetic solenoid used in machine setup image



(b) Media characterization (a) FTIR, (b) SEM and TEM

Fig. 3 **a** Electromagnetic solenoid used in machine setup (**b**) Media characterization **a** FTIR, **b** SEM and TEM image

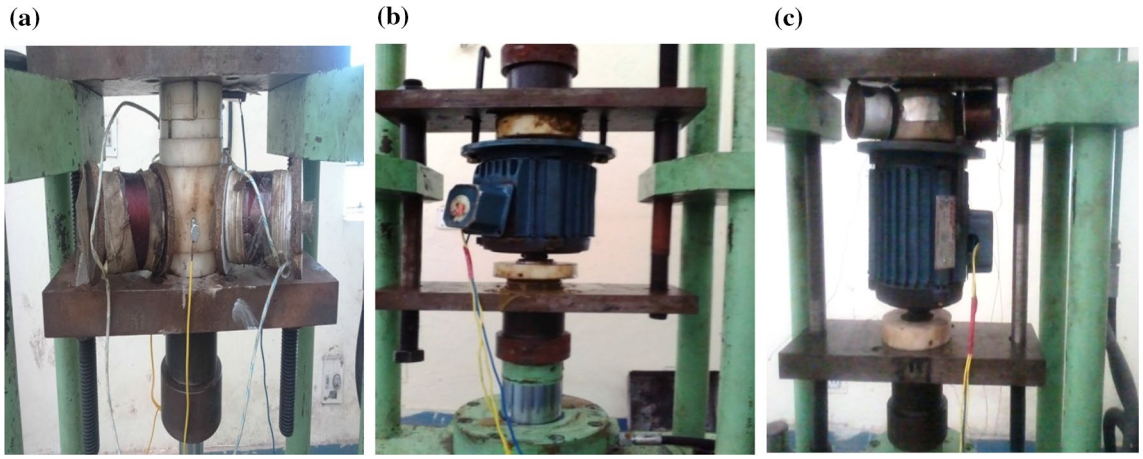


Fig. 4 AFM setup portion between upper and lower machine cylinder (**a**) Magnetic, electrolytic action, **b** rotational, electrolytic action, **c** Magnetic, rotational action

in which the electromagnetic solenoid location is shown clearly. It can be easily understood that the upper and lower piston-cylinder arrangements, i.e., machine and media cylinders are supplied hydraulic pressure from the machine hydraulic power pack. Further, the electrochemical and magnetic setup power supplies are also shown in addition to extrusion pressure cylinders arrangement in the hybrid AFM setup.

2.3 Experimental Technique Used

First of all, best media selection was done using response surface methodology (RSM), then Taguchi L9 OA method was applied for experimentation purpose so that best parameter selection can be found out to optimize the results. Then different types of workpiece to be finished were selected and different types of abrasive particles that result in finishing of the material from the work surface and their mesh size and their mixing ratio were selected. Many researchers have also arrived at the best optimum condition of parameters selction in their study in the past. In AFM of slit fabricated by Electric Discharge Machining (EDM) process, the optimum parameters were 150 µm, concentration of 50%, pressure 6.7 MPa and time of machining was 30 min [36]. Hence, parameters selction is very important part of experimentation. To make the conventional machine hybrid, some attachments were added that contributed to more number of input parameters i.e., rotational speed of workpiece which helps in proper intermixing of media inside the workpiece, electrolytic voltage that helps in increasing more material removal using electrochemical machining action. In addition to analyse the effect of magnetic field, the magnetic flux parameter of the electromagnetic solenoid was also considered. The complete list of parameters and their levels taken is given in Table 1.

The experimentation work was carried out in a very systematic manner, in which firstly the best polymer media was selected in a scientific manner by performing experiments on five different types of media one by one taking the same three input variable parameters in each case. Then further the experiments were performed on hybrid ECMR-AFM setup by using the best media selected from the previous experimentation. A surface model was prepared using RSM technique taking into consideration four input parameters i.e., type of media, extrusion pressure, volume of media and number of cycles. Then, the comparative study was performed between the material removal results obtained from the RSM technique and that from Taguchi L9 OA technique. In the end, the comparison of machine, conventional AFM and hybrid ECMR-AFM parameters were compared to check whether the newly developed machine setup is more efficient in terms of material removal and surface finish or not. Lastly the Matlab fuzzy logic optimization technique and GRA was performed on the developed analogy of input parameters to obtain optimized output results to obtain the best combination of input parameters and then comparative results obtained from other techniques were analysed. The aim of experimental analysis is to prove the theoretical results in a practical way that can be used in industries.

3 Results and Discussion

The results and discussion section is the main output of the experimental and analytical research carried out by the researchers in the present study. It comprises the output

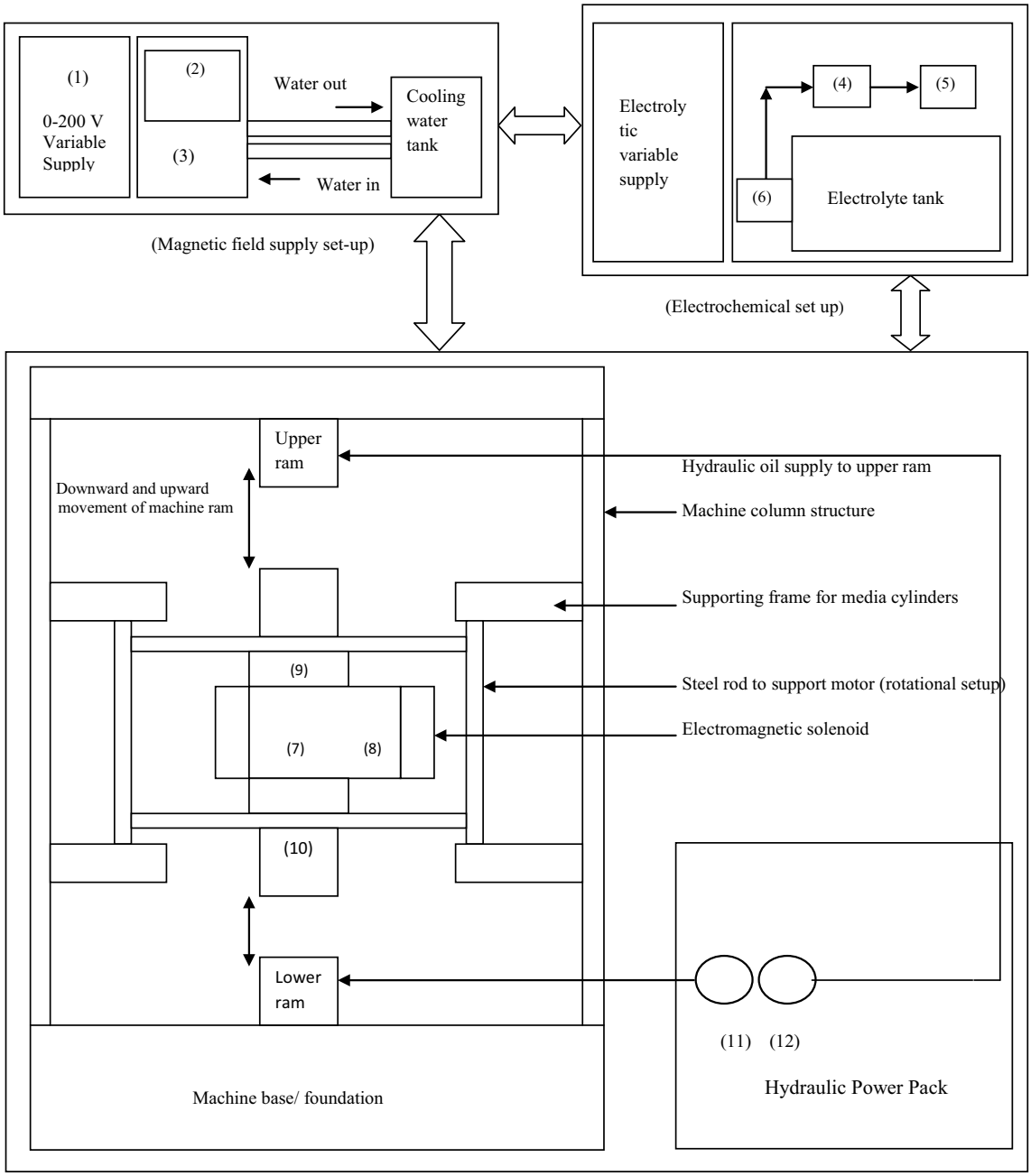


Fig. 5 AFM Machine set-up block diagram: 1. Transformer, 2. Electromagnet, 3. Transformer oil tank, 4. Flow-sensor, 5. Pressure gauge, 6. Motorpump, 7. 3-Phase motor, 8. Speed controller, 9. Fixture, 10. Lower media cylinder ram, 11. Lower machine cylinder pressure gauge and controller, 12. Upper machine cylinder pressure gauge and controller

results, i.e., material removal and roughness behavior with respect to input parameters taken on the hybrid ECMR-AFM setup. Further the mathematical and computational modelling of the process alongwith Taguchi L9 OA, fuzzy logic, GRA optimization technique results are also discussed in details.

3.1 Effect of Input Parameters on Material Removal

In this section, the output material removal (MR) results obtained during the abrasive flow machining process is analysed considering the effect of input parameters taken. Firstly, the different polymer media that were prepared were utilized in the process and their effect was calculated. In abrasive electrochemical grinding (AECG), the electrochemical grinding of sintered carbides G20, titanium alloys

Table 1 Input parameters used and their levels

S.No	Input parameters	Unit	Level-1	Level-2	Level-3
1	Electrolytic voltage (ev)	Volt	6	12	18
2	No. of cycles (nc)	N	2	4	6
3	Pressure (ep)	bar	10	20	30
4	Speed of rotation (rs)	RPM	120	140	160
5	Workpiece type (wt)	–	Al	Brass	C.I
6	Abrasive particle type (at)	–	Silicon carbide	Alumina	Mixture of both
7	Magnetic flux (mv) (corresponding voltage)	tesla (V)	0.5 (50)	0.75 (125)	1.0 (225)
8	Abrasive Ratio (ar)	–	2:1	1:1	1:2
9	Mesh size (am)	µm	100	200	300

Table 2 Effect of different polymer media used on material removal

Exp. No	Material removal (MR) (mg)				
	SBR	NTR	SR	PBS	NR
1	3.6	3.6	3.9	2.5	2.3
2	4.4	3.8	3.8	2.1	2.3
3	4.0	3.1	3.9	1.7	2.3
4	3.0	3.7	3.5	3.2	3.1
5	4.9	2.5	3.3	2.4	2.4
6	3.4	2.5	3.1	2.1	3.8
7	3.6	2.3	4.0	2.9	3.33
8	4.9	3.8	3.2	2.0	3.0
9	3.5	3.1	3.8	2.6	3.3
Avg	3.8	3.1	3.6	2.3	2.8

WT31 was done and it was concluded that wear in conventional grinding was 6 times more as compared to AECG, as investigated by Zaborski et al. [37]. Hence, the role of abrasive particles is very high in material removal from the work surface. The experiments were performed based on Taguchi L9 OA, by taking first three parameters, i.e., electrolytic voltage, number of cycles and pressure as variables and rest parameters as constant, as shown in Table 1. The five different types of prepared media are used to check the best usable media in AFM process. The Table 2 consists values of MR during AFM process in case of different media viz. SBR, NTR, SR, NR and PBS, used. The MR was found to be minimum in PBS, i.e., 2.39 mg, maximum in SBR, i.e., 3.88 mg and intermediate results were obtained in NTR,

Table 3 ANOVA table for material removal

Source	Sum of squares	DOF	Mean square	F-value	*P*-value Prob. > F	
Model	10.09	14	0.72	5.68	0.0009	Significant
Media type (mt)	0.41	1	0.41	3.27	0.0906	
Extrusion pressure (ep)	6.023E-003	1	6.023E-003	0.047	0.8304	
No. of cycle (nc)	0.029	1	0.029	0.23	0.6372	
Vol. of media (vm)	0.90	1	0.90	7.07	0.0179	
mt-ep	4.977E-003	1	4.977E-003	0.039	0.8456	
mt-nc	0.058	1	0.058	0.46	0.5092	
mt-vm	0.072	1	0.072	0.56	0.4641	
ep-nc	0.15	1	0.15	1.20	0.2899	
ep-vm	0.026	1	0.026	0.21	0.6544	
nc-vm	0.073	1	0.073	0.58	0.4596	
$(mt)^2$	3.30	1	3.30	25.98	0.0001	
$(ep)^2$	4.67	1	4.67	36.83	< 0.0001	
$(nc)^2$	2.52	1	2.52	19.84	0.0005	
$(vm)^2$	0.062	1	0.062	0.49	0.4937	
Residual	1.90	15	0.13			
Lack of fit	1.54	10	0.15	2.14	0.2075	Not significant
Pure error	0.36	5	0.072			
Cor total	12.00	29				

NR and SR media, as given in Table 3. It can be concluded that maximum MR occurs in case of SBR media usage, i.e., 4.9 mg at electrolytic voltage 18 V, number of cycles 4 and extrusion pressure 10 bar. But when voltage taken was 6 V, number of cycle 6 and pressure 30 bar, MR was found to be minimum, i.e., 1.7 mg in case of PBS media usage. The reason may be the poor flow ability of PBS and due to lesser input voltage. Hence, it is proven that SBR results in more removal of material as compared to PBS, further, electrolytic voltage increases the MR, whereas pressure and number of cycles effect on MR is lesser as compared to that of voltage.

In Table 3, Analysis of Variance (ANOVA) technique applied shows that model is significant including media, pressure, number of cycles and volume of media, denoted by letters A,B,C, and D, respectively. As seen from Table 3, which is based on analysis of variance for the workpiece material removal, clearly indicates that the model is significant and the lack of fit is found to be insignificant. The *P*-values of the parameters, i.e., A, B, C and D are also significant alongwith their combination sets, i.e., AB, BC, etc., and their square values, the model degree of freedom being 14. The input parameters taken were in 4 levels, i.e. A (1. NR, 2. NTR, 3. SBR, 4. PBS), B (1. 15, 2. 20, 3. 25, 4. 30), C (1. 6, 2. 7, 3. 8, 4. 9) and D (1. 200, 2. 215, 3. 230, 4. 245), the units of A, B, C and D being dimensionless, bar, number and ml respectively.

The surface plots obtained showing output material removal results in the RSM technique are as shown in Fig. 6. the input variables, i.e., mt, ep, nc and vm are taken on horizontal plane, i.e. X1 and X2 axis, while output variable, i.e.,

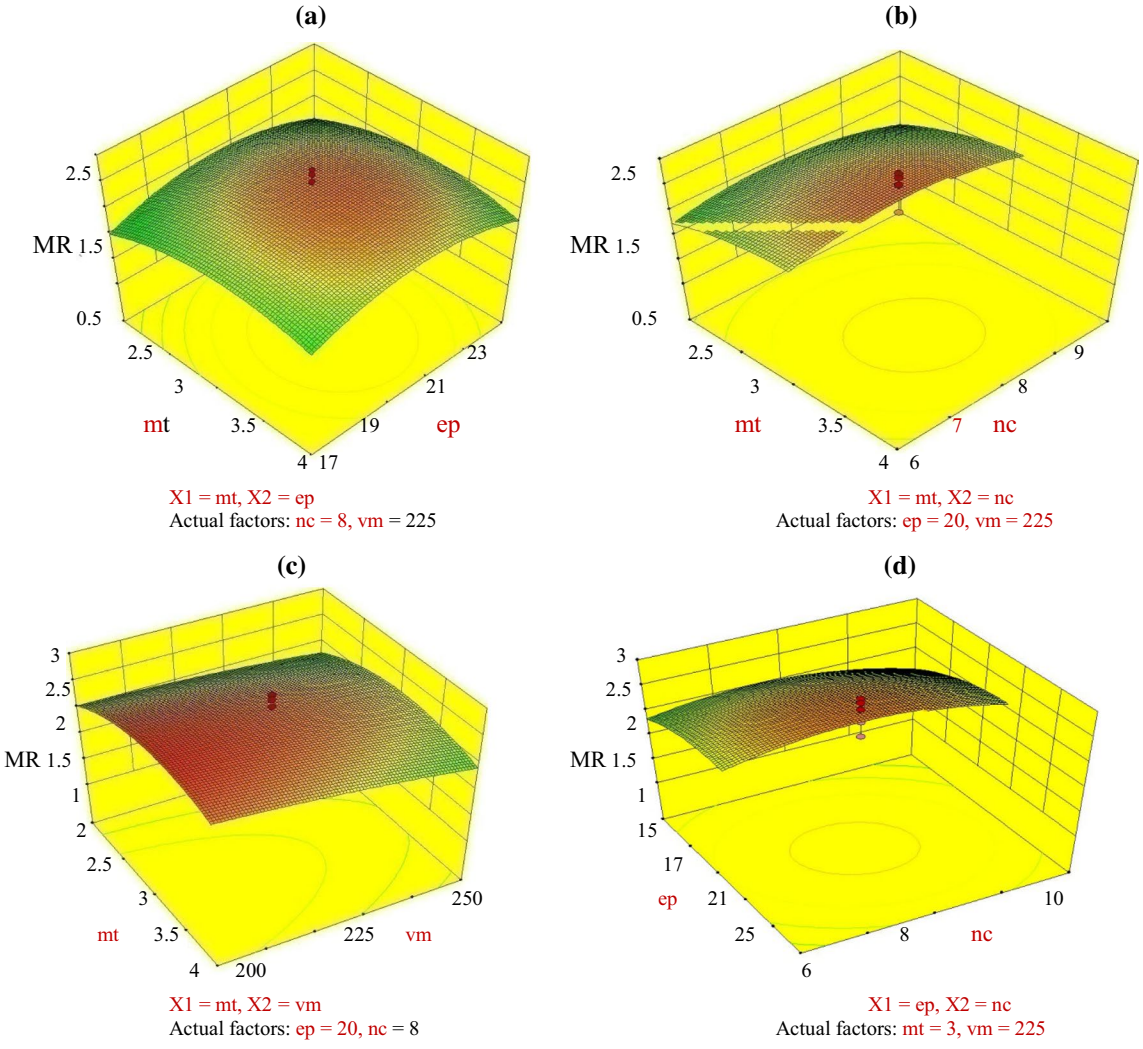


Fig. 6 3D Surface model (**a**) Type of media (surface view from top), **b** Type of media (surface view from front), **c** Extrusion pressure (surface view from top), **d** Extrusion pressure (surface view from front)

MR is taken on Y axis. The Fig. 6a and b shows top and front view respectively, that are based on constant type of media parameter used in which the input parameters combination set mt-ep and mt-nc was used respectively. It can be inferred from the Fig. 6a and b that MR value lies between 1.5 mg to 1.9 mg, maximum being at extrusion pressure 21 bar and number of cycles 7. In Fig. 6c and d, the output MR results are based on constant extrusion pressure used in which top and front view surface plots are plotted to analyse the MR results based on BC and AD input parameters combination sets. The hybrid electromagneto rotational abrasive flow machining (ECMR-AFM) setup resulted in higher MR and better surface integrity. For calculating the effect of magnetic field on the hybrid AFM set-up, the three input parameter i.e. magnetic voltage or flux, ECM rod size and voltage were taken and it was found that both electric and magnetic lines of force caused increased velocity of particle hitting the work surface. The rod size first increases the metal removal and then its effect is negligible. The Taguchi L-9 OA is applied to optimise the result in which 9 experiments have been performed using different combination of 3 input parameters, the results shown in Table 4.

Abrasive type (at), abrasive mesh size (am), abrasive ratio (ar), electric voltage (ev), rotational speed (rs), magnetic voltage (mv), workpiece type (wt), number of cylces (nc), extrusion pressure (ep) i.e. machine parameters used are shown in Table 4. As shown in Table 4, the combination of input parameters for experiments using Taguchi L9 OA technique, is made in a scientific manner so as to understand the overall effect of abrasive particles, voltage and workpiece used, etc. The first combination, i.e., (at, am, ar), signifies the importance of abrasive particle, its size and the mixing ration in the machining efficiency. The second combination (ev, rs, mv) is based on the hybrid ECMR-AFM machine setup fabricated, so as to analyse the effect of different types of forces on the output finishing results. The third combination, i.e., (wt, ep, nc) gives us the data of conventional machine parameters effect on MR and Ra results.

It can be inferred from output results, as shown in Table 4, that the second combination set results in overall increase in material removal and percentage improvement in Ra value as compared to conventional AFM machine used for experimental purpose. Hence the hybrid machine setup used by the authors resulted in 34.5% and 17.8% rise in % Ra improvement and MR respectively, as compared to traditional setup, the average output results data is as shown in Table 4.

The applied magnetic field at starting increases material removal at slower rate but afterwards its increase is high, as shown in Fig. 7a. In Fig. 7b, it is clear that the MR rate value increases from 0.2 mm^3/min to 3.0 mm^3/min when ECM rod size of diameter 6 mm is put inside the workpiece instead of rod of 4 mm diameter size. This indicates that by increasing the diameter or decreasing the gap between anode and cathode from 6 to 4 mm, the more MR occurs. Afterwards, using thick thick ECM rods greater than 6 mm results in very less improvement in removal of material. The approximately same trend is shown in Fig. 7c, in case of electrolytic voltage, in which MR increases at fast rate when electrolytic voltage increases from 6 to 12 V, but further increase in voltage increases MR at slower rate.

The ANOVA calculations for S/N ratio is as shown in Table 5. The three experimental runs of material removal of each workpiece is shown alongwith mean, variance and sum of squares. It is found that all the three input parameters taken are significant, their degree of freedom being 2 each.

The material removal effect with and without rotational effect is clearly depicted in Fig. 8a in which 1600 rpm speed resulted in maximum removal of material. When rotational force in not employed in AFM process, then on an average 2.9 mg material gets removed, while using rotational setup results in removal of material upto 4.5 mg. As shown in Fig. 8b, it is clear that when speed of rotation is increased from 800 to 1600 rpm, the MR value drastically increased from 2.5 µm to 4.5 µm. The power consumption of the machine vs volume of material removal in case of different polymer media developed is shown in Fig. 8b, where it was

Table 4 Material removal and roughness output values

Parameters	at, am, ar		ev, rs, mv		wt, ep, nc	
Exp. No	MR	Ra%	MR	Ra%	MR	Ra%
1	5.8	24.01	7.8	34.01	6.8	24.4
2	5.03	16.4	7.3	26.4	6.03	17.9
3	4.86	30	6.9	30	5.86	30.9
4	6.64	18.8	6.9	28.8	5.64	18.2
5	9.19	14.33	9.2	24.33	8.19	14.69
6	8.14	13.1	9.5	23.1	8.14	13.98
7	7.1	25	10.1	35	8.1	25.91
8	6.5	8.9	10.5	9.9	8..5	8.9
9	10.5	10.9	12.9	10.9	11.5	10.99
Avg	7.06	18.1	9.02	24.75	7.64	18.41

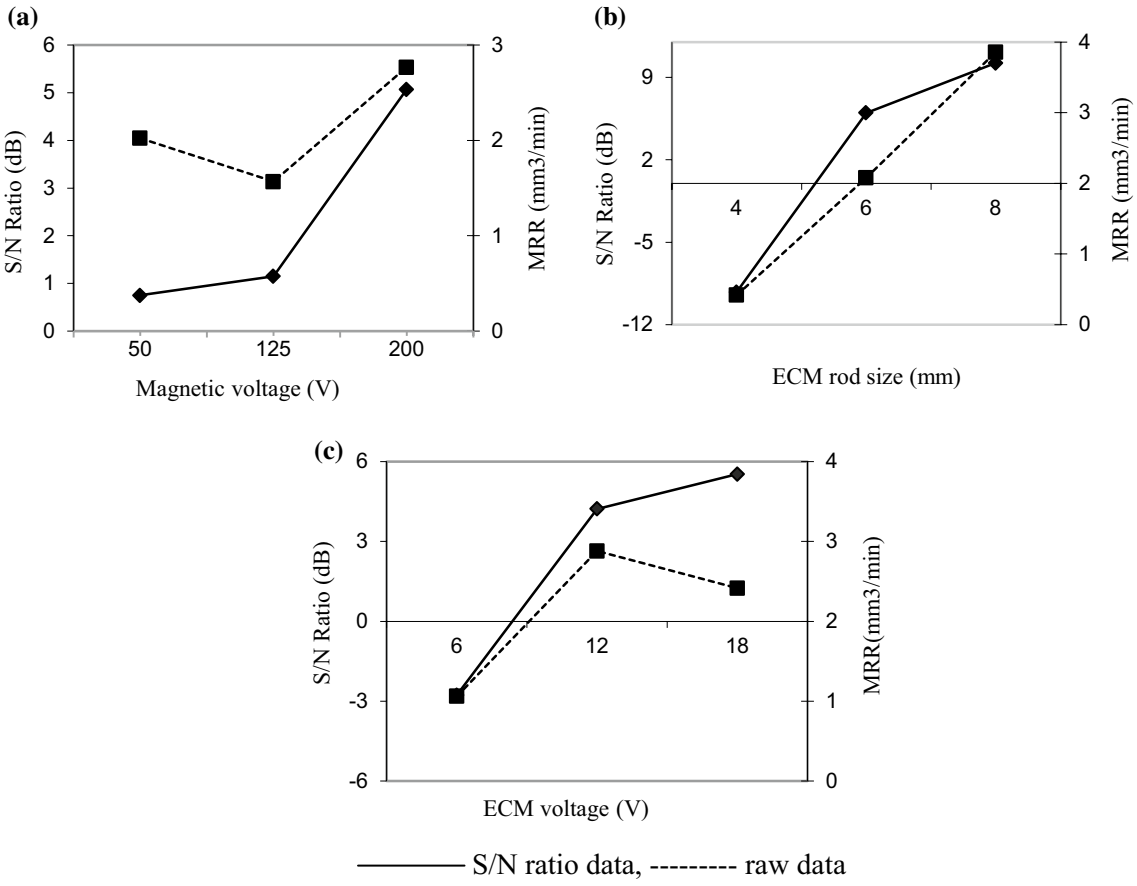


Fig. 7 S/N ratio and raw data output graphs of material removal rate vs (**a**) Magnetic voltage, **b** ECM rod size, **c** ECM voltage respectively

Table 5 ANOVA table

Source	Sum of square	Degrees of freedom	P	V	F-value	Pool
Electrolytic rod	630.33	2	81.24	80.50	19	No(significant)
ECM voltage	120.44	2	16.14	300.12	19	No(significant)
Magnetic flux	33.22	2	4.46	15.54	19	No(significant)
Error	2.03	2				No(significant)

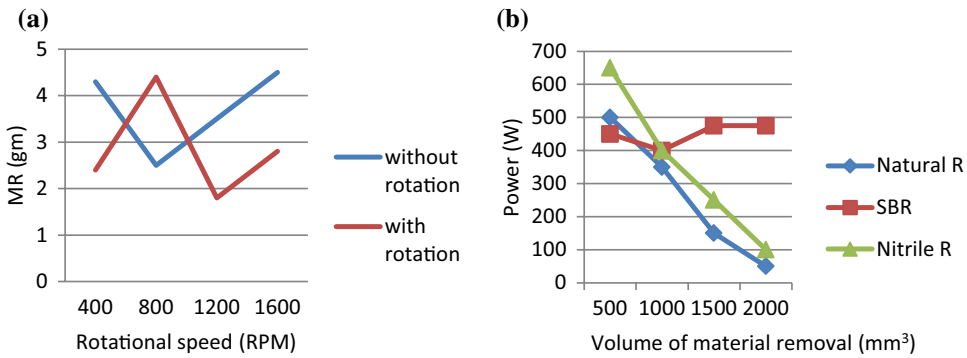


Fig. 8 **a** MR vs rotational speed plot, **b** Power vs volume of material removal plot

found that natural rubber based media required more power as compared to nitrile and SBR for same volume of material removed. The power consumption in ECMR-AFM process got reduced from 0.65 kW to 0.1 kW when NR media was used, while SBR media usage resulted in constant power consumption of about 0.5 kW with increase in material removal of the workpiece. The material removal volume was calculated by finding the difference in weights of the workpiece before and after machining process.

The proper validation has been done for both experimental and analytical output results; so that we can obtain best optimized result i.e. higher material removal and lower surface roughness value. In Fig. 9a, material removal effect on magnetic voltage and number of extrusion cycles is explained. The increase in number of cycles from 4 to 8 gradually increased the MR as seen from the surface plot and material removal first increased and then decreased by increasing the magnetic voltage. The analytical methods, ie., surface and interval plots were applied so that the complete overview of the various input parameters' effect on output responses can be understood. In the interval plot as shown in in Fig. 9b, the 3 intervals of rod size taken at 4, 5, 6 mm, the MR results obtained are about 0.17 gm at 4 and 5 mm rod size, while 0.125 gm at 6 mm rod size. From this data, we can easily infer that the MR increases when we use ECM rod of size more than 4 mm and the same inference was drawn in the previous section. Hence, it is concluded that in both cases, i.e., Taguchi L9 OA output results and the interval plots, the results are nearly same.

3.2 Effect of Input Parameters on Surface Roughness

In this section, the surface roughness improvement is analysed and the effect of various input parameters is understood in details. Various authors have obtained significant results in surface roughness reduction. Firstly, their brief work is introduced in the section and then the output results obtained by the present author in doing experiments on ECMR-AFM setup are discussed in detail. The electrochemical finishing/polishing (ECF/ECP) and its variants helped to achieve micro to nano range of surface finish, the parameters affecting the output results included electrolyte properties, position and motion of electrode [38]. If EDM and ECM lapping was used, then surface roughness was improved upto 0.07 mm in 2 min [39].

In double disc MAF of paramagnetic materials, Al disc with 4 NdBFe magnets is used. Al is used as casing because permeability is one, hence magnetic lines of force are not affected [40]. In case of finishing of workpiece material, the utmost requirement is that surface obtained should be of very fine finish and integrity. Hence, the input parameters are controlled in such a wayso as to get nano-level fine finished products. These products are used in precision instruments where high degree of accuracy is needed, in medical and surgical dentures where medical care has to be done in a very safe environment since the life and health of an individual is considered. The effect of finishing time of workpiece, extrusion cycles, machining pressure on the output surface roughness response has been analysed graphically in this section. The matrix of mold injection surface of initial Ra value 0.65 μm was polished till Ra reached 0.07 μm in 1500 s in AFM [36]. In the prepared machine setup, the

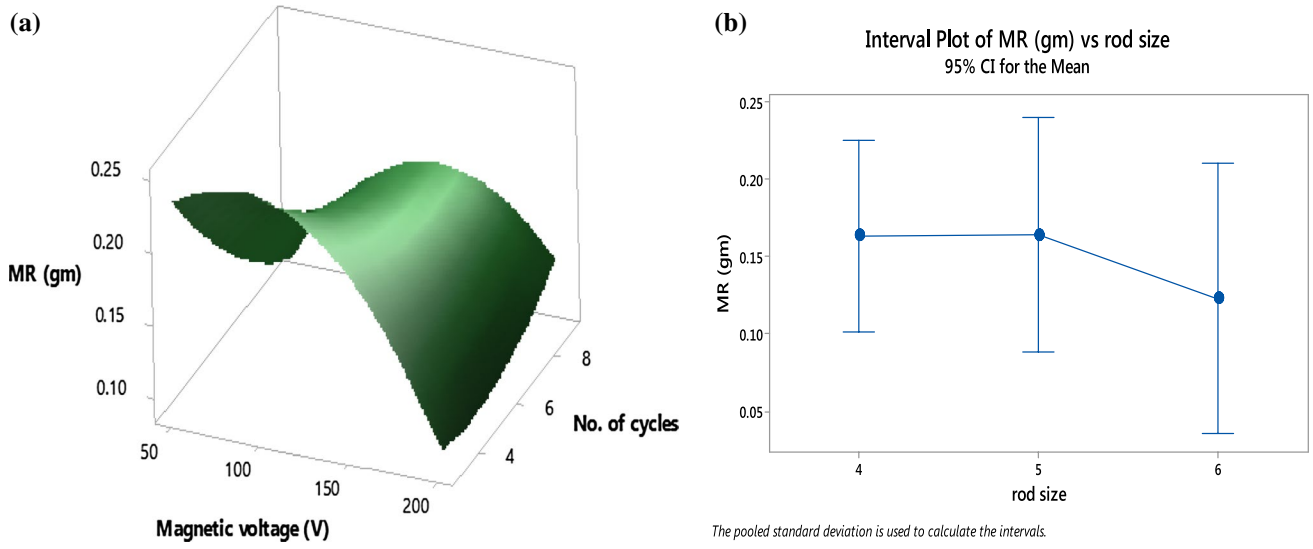


Fig. 9 **a** Surface plot of MR vs magnetic voltage, **b** ANOVA interval plot of MR vs rode size

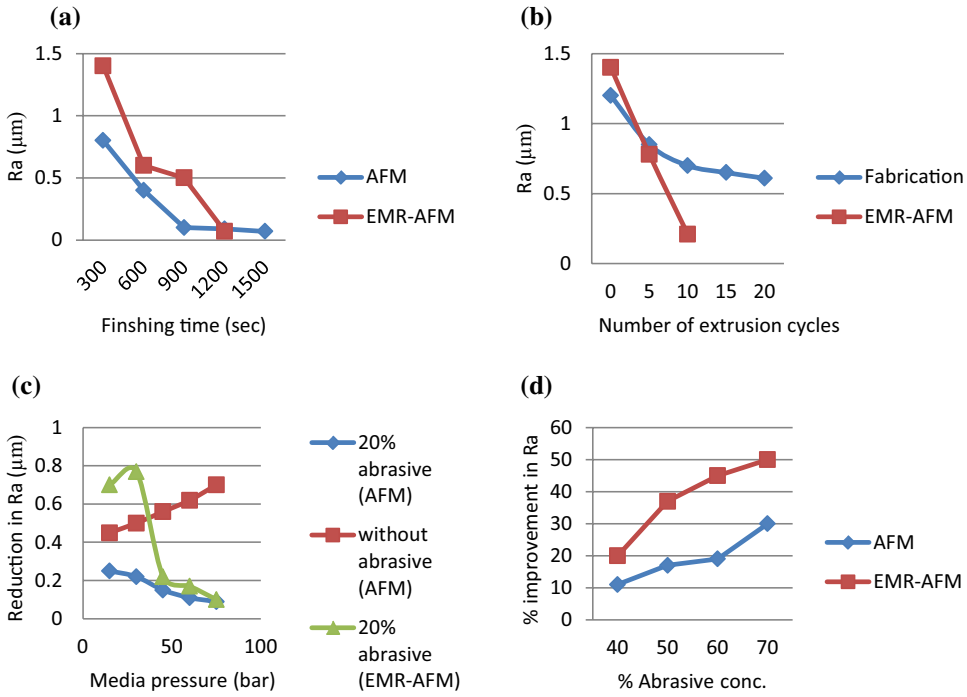


Fig.10 **a** Reduction in Ra vs (a) finishing time [27], **b** extrusion cycles, **c** media pressure, **d** % abrasive concentration plot [36]

surface roughness was reduced from initial Ra value of 1.4 to 0.07 µm in 1200 s, i.e. 40% more roughness reduction in 20% lesser time as compared to conventional AFM process, comparison shown in Fig. 10a. The percentage improvement in surface roughness was obtained and compared with the input parameters viz. vibration amplitude, pressure and time of machining the temperature and viscosity of media taken was 300 K and 730 poise respectively as shown in [41], while initial Ra kept between 1.2–1.4 µm and concentration of abrasive in media by weight was 60:40. The surface roughness value was reduced from 0.8 µm to 0.2 µm in rotary ultrasonic machining [42], the surface roughness value was drastically reduced from 1.4 µm to 0.11 µm in the present working set up, as shown in Fig. 10b. The roughness value decrease with number of extrusion cycles as shown. It can be inferred from Fig. 10b, that the number of machine extrusion cycles requirement is decreased from 15 to 8 in the present hybrid setup utilized in experimentation. The conventional setup resulted in roughness reduction from 1.25 μm to 0.6 μm, while the hybrid setup decreased Ra from 1.5 μm to 0.2 μm, i.e., 53.8% more improvement in roughness. The main aim is to achieve high material removal, better surface finish, reduced sub-surface damage. In Fig. 10c and d, the surface roughness percentage improvement is more in hybrid electrolytic magneto rotational AFM setup as compared to conventional AFM process. In case of conventional AFM process, when 20% abrasive particles and when no abrasive particle is used, then Ra value get reduced from 0.3 µm to 0.12 µm at gradual rate with increase of media pressure and roughness value increased with increase of pressure respectively. As shown in Fig. 10c, it is clear that in ECMR-AFM setup, the roughness value got reduced from 0.7 µm to 0.19 µm at 50 bar extrusion pressure value. Hence it is concluded that more decrease in surface roughness can be achieved in hybrid machine setup at half the extrusion pressure value. It can be inferred from Fig. 10d, that 10% more improvement in roughness value results in case of hybrid AFM setup as compared to conventional AFM process when percentage abrasive concentration is increased in polymer media. This may be due to increase in number of cutting edges of abrasive particles participating in machining process.

3.3 Matlab Fuzzy Logic and Grey Relational Optimization of Material Removal and Surface Roughness Results

Grey relational analysis (GRA) was applied to study and analyse the output results obtained in ECM process by Liu et al. [43]. GRA optimization technique was applied by Lin and Lee [44]. in magnetic assisted EDM of SKD 61 workpiece Taguchi L18 OA technique was applied and parameters taken were machining polarity, high voltage auxiliary current (IH), servo reference voltage (SV). After getting the output optimized results of MR and Ra values in Taguchi L9 OA and RSM technique, the results are optimized in multi objective fuzzy logic optimization and grey relational analysis and then compared and to get to a conclusion that which technique is best. In fuzzy logic

Table 6 GRC and MPCI rank values corresponding to Taguchi L9 OA exprimentation

Exp. No	% Ra	Normalised (N)	Deviation (D)	GRC	MR	N	D	GRC	MPCI	Rank
1 (111)	20.37	0.09	0.81	0.38	4.2	0.00	1.00	0.33	0.06	9
2 (122)	16.4	0.00	1.00	0.33	10.8	0.43	0.57	0.46	0.25	8
3 (133)	50.8	0.75	0.25	0.66	12.4	0.53	0.47	0.51	0.55	3
4 (212)	30.3	0.30	0.70	0.41	14.4	0.66	0.34	0.59	0.42	6
5 (223)	38.1	0.47	0.53	0.48	19.6	1.00	0.00	1.00	0.65	1
6 (231)	62.4	1.00	0.00	1.00	10.2	0.39	0.61	0.45	0.64	2
7 (313)	36.5	0.44	0.56	0.47	7.2	0.19	0.81	0.38	0.31	7
8 (321)	50.4	0.74	0.36	0.58	8.2	0.26	0.74	0.40	0.45	5
9 (332)	53.9	0.82	0.28	0.64	9.8	0.36	0.64	0.43	0.54	4

Table 7 Shear rate, pressure, velocity values at different locations during media flow

S. No	X (m)	Y (m)	Z (m)	Velocity V_Z (m/s)	Shear rate (1/s)	Pressure (Pa)	Shear stress (Pa)	Temp. of fluid (K)
1	0.00069	0.0657	0.0046	1.93	2037	11.76	0.04	293.19
2	0.00069	0.0636	0.0046	2.00	1168	11.73	0.06	293.29
3	0.00070	0.0657	0.0046	1.87	1957	11.71	0.04	293.19
4	0.00070	0.0636	0.0046	1.90	1057	11.70	0.06	293.28
5	0.00067	0.0615	0.0046	2.00	1578	11.41	0.07	293.40
6	0.00067	0.0594	0.0046	1.99	2035	11.10	0.09	293.54
7	0.00070	0.0615	0.0046	1.88	1478	11.39	0.07	293.40
8	0.00070	0.0594	0.0046	1.86	1958	11.08	0.09	293.54
9	0.00942	0.0501	0.0094	0.00	11.63	67.53	0.07	294.06
10	0.00702	0.0501	0.0114	0.00	6.155	67.48	0.03	293.76

optimization process, first of all the membership functions were defined i.e. very low to very high in 5 levels, shown in Fig. 8.

Highest Multi Performance Characteristics Index (MPCI) value corresponds to lowest rank as shown in Table 6. In GRA, firstly, the normalization of data is done, then for various experiments the grey relational coefficients are calculated, as shown in Table 7. These values range from 0 to 1. At last grey relational grade is calculated. The Eq. (1) is used for calculating GRC value.

$$\mathrm{GRC} = \frac{d_{\min} + e.d_{\max}}{d(k) + e.d_{\max}} \tag{1}$$

where d is the value corresponding to the experiment run ranging from k = 1–9, e is value taken corresponding to the number of output responses taken. The three input variables taken were magnetic flux, electrolytic voltage and rotational speed and each parameter has 3 levels. As it is clear from Table 6, rank 1 corresponds to the condition of magnetic flux of 0.75 Tesla corresponding to 125 V magnetic voltage, 12 V electrochemical voltage and 160 RPM rotational speed. In Table 7, the final output results in terms of MR and Ra are listed for both fuzzy logic and grey relational analysis. The GRC values range from 0.0 to 1.0 in case of material removal of 4.2 µm to 19.6 µm respectively, while it range from 0.33 to 1.0 for percentage roughness improvement value of 16.4% to 62.4% respectively. Hence, it is concluded that GRC value increases if material removal increase and percentage improvement of roughness increases. After Grey Relational Coefficient (GRC) calculations, the deviation is found out and then at last MPCI value is calculated from Matlab fuzzy logic software. The MPCI values ranges from minimum 0.06 to maximum 0.64, i.e., it signifies the best combination of input parameters so as to get intermediate value of material removal and improvement of roughness value, i.e. optimized combined result, hence called multi objective optimization process.

As shown in Fig. 11, the Fuzzy Inference System (FIS) variables are set as output variables of finishing operation and corresponding MPCI variables are generated. It is also shown that membership function plots are also set according to the input parameters, as low (L), very low (VL), medium (M), large (L) and very large (VL) based on the effect of input parameters, i.e., magnetic flux, electrolytic voltage and rotational speed.

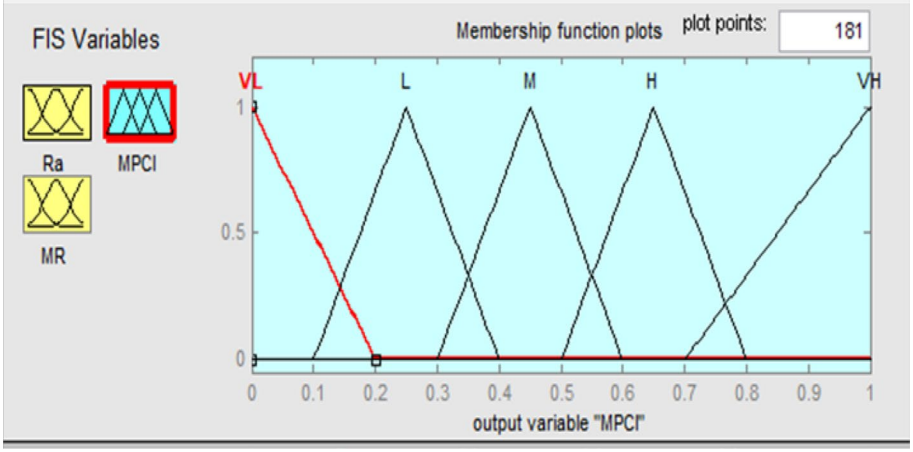


Fig. 11 Fuzzy logic operators alongwith membership plot and FIS variables

4 Mathematical and Computational Modelling of Ecmr-Afm Process

The modelling of the hybrid ECMR-AFM process is done which includes voltage, cutting forces and magnetic flux calculations in elecrochemo magnetic AFM process, analysis of flow of media, i.e., shear rate, velocity, pressure calculations, etc., the role of computer neural network and algorithms in AFM process is analysed. Further, the computational flow analysis of the flow of media through the newly designed and fabricated ECMR-AFM machine fixture is done using Ansys software.

4.1 Role of Computer Optimization Software and Algorithm in AFM Process Modelling

The demands of cost feasible products led to the development of advanced non conventional finishing processes like abrasive flow finishing process. In AFM process, abrasive media is extruded across the surface to be finished in single direction or two-way or orbital [45]. Chryssolouris and Guillot [46] modelled a machining process using neural network and multi-regression method. The neural network approach has been extensively used in abrasive flow machining. Jain et al. [47] developed the equations for material removal rate (MRR) and surface roughness

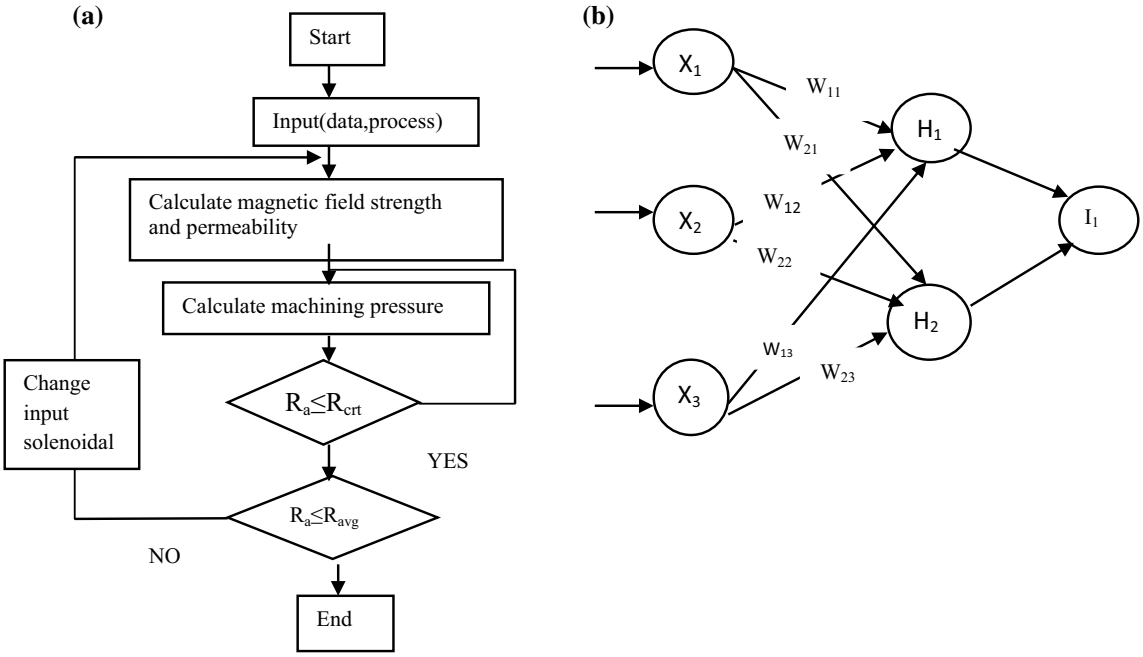


Fig. 12 **a** Flow-chart representing magnetic forces effect, **b** Neural network approach in AFM Silva and Geough [43] have done a lot of work on computer applications in unconventional machining

(Ra) by multi variable regression and corresponding coefficients were evaluated using least squares method. The back propagation neural network is shown in Fig. 12b. Petri et al. [48] used the neural network model to predict the surface finish in abrasive flow machining. The developed to set machine parameters of AFM, by pairing with meta-heuristics, which assisted neural network model in extrapolating. The response of node of hidden layer,

$$h_j = f\sum W_{ji}X_i \tag{2}$$

for output layer,

$$r_i = f\sum v_{ij}h_j \tag{3}$$

where w is weight of connection to hidden layer as shown in Eqs. (2) and (3). The function f is transfer function given by Eq. (4).

$$f(t) = \left(1 + e^{-t}\right)^{-1} \tag{4}$$

Shinmura and Aizawa [49] developed a simulation to study on magnetic assisted abrasive finishing process. After analysis of the results of magnetic finishing the simulation results showed that the magnetic flux between north and south poles is maximum when the air gap between the poles is minimum. Figure 12a clearly explains the flow chart and step by step simulation of magnetic assisted finishing process. Jain and Jain [50] simulated the surface profile in abrasive flow machining by assuming the random distribution of abrasive grains in media. Uhlmann et al. [51] applied Maxwell model to polymer media by Computational Flow Dynamics (CFD) simulation of AFM process.

Relaxation time,

$$t = z/c \tag{5}$$

$$C'(w) = C_0 + \sum C_k.\left(wt_k\right)^2/\left[1 + \left(wt_k\right)^2\right] \tag{6}$$

The relaxation time calculation is given in Eq. (5), while spring stiffness calculation for next series is done in Eq. (6). Jain and Jain [52] used back-propagation neural network model to optimize abrasive flow machining process.

4.2 Voltage and Magnetic Flux Calculations

A model was developed in which electrode gap was simulated and modelled and final micro tool diameter was predicted within the accuracy of 94% or 6% variation range to develop tungsten tools by Mathew et al. [54]. In magnetic field assisted AFM process, the active abrasive particles actual participating in cutting action were calculated using developed model by Das et al. [55]. The laminar Newtonion fluid was utilized in the process and model was simulated at variable volume fraction of media using Ansys software to study forces and analyse the media flow in AFM process [56]. Nearly 56% and 49% improvement in Ra in aluminium and mild steel workpieces respectively using flyash besed polymer media has been achieved in AFM process [57]. The surface roughness improvement has been achieved by revealing over and under polishing using roughness model developed [58]. The calculation of magnetic force and the removal of material is the prime focus which resulted in the need of a mathematical model. Hence model was generated by assuming a magnetic particle of radius r and mass m and is of spherical shape. We apply continuity Eq. (7), based on mass conservation principle.

$$\left(\frac{4}{3}\pi r^3\right) = 1/\rho \tag{7}$$

differentiating we get $\frac{\partial}{\partial m}\left(\frac{4}{3}\pi r^3\delta r\right) = -\frac{\delta\rho}{\rho^2}$.where ρ is the density of magnetic particles. The abrasive particle equation of motion is given by Eq. (8). In Eq. (9), the pressure calculation is shown.

$$\frac{\partial^2 r}{\partial t^2} = -2\pi r\frac{\partial P}{\partial m} - 2\frac{Gm}{r} \tag{8}$$

The total pressure

$$P = p + \frac{H^2}{8\pi} \tag{9}$$

Differentiating the above equation, we get the solution from Eq. (10),

$$\delta P = \delta p + \frac{H\delta H}{4\pi}$$

$$\frac{\partial^2 r}{\partial t^2} = -2\pi r\frac{\partial}{\partial m}\delta P + 4\frac{Gm}{r^2}\delta r \tag{10}$$

The magnetic field B is calculated as in Eq. (11).

$$B = \frac{\mu_o}{2}\frac{NIx^2}{\left(x^2 + z^2\right)^{3/2}} \tag{11}$$

$= 10$ Gauss.

The magnetic flux is calculated as given in Eq. (12).

$$\varphi = AB\cos\alpha \tag{12}$$

$= \pi \times 4.5^2 \times 10 \times \cos 60^o = 308$ G/A.

According to Faraday's law of induction as shown in Eq. (13),

$$V = -N\frac{d\varphi}{dt} \tag{13}$$

$$= -3200 \times 4.5^2 \times \cos 60^{\circ} \times 10/120 = 2784 \text{ V}$$

$$B = 0.5\, B_P \cos \omega t \tag{14}$$

$$\omega = \frac{10}{0.5 \times 40000} = 0.75 \text{ rad/s}$$

where ω is angular frequency of alternating current., which is calculated from Eq. (14).

Hence

$$V = 0.5\, \omega N A B_P \cos \alpha .\, \sin \omega t \tag{15}$$

$$V = M \frac{di}{dt} \tag{16}$$

The Eqs. (15) and (16) are used for calculating voltage using current gradient.

The erosion is given by Eq. (17), which denotes the rate of removal of metal from the work surface.

$$R = \frac{dm}{dA_m} \tag{17}$$

$$M = k \rho_w V_A N \tag{18}$$

In Eq. (18), mutual induction expression is shown.

Figure 13 shows a schematic drawing in which an active particle is subjected to a number of external forces. The direction of current is shown by fingers curling orientation of the right hand. It is seen that the different kinds of forces viz. magnetic, electrolytic and centrifugal force act on the abrasive particle. The piston cylinder arrangement above and below the fixture containing workpiece provide the required driving force to the media. This gives extrusion pressure to the abrasive laden polymer media which results in material removal from the surface of the job. The different forces incorporated in the system are as shown in Fig. 14a, i.e., driving force transferred by the pressure of the media and normal force mainly produced by cylinder total pressure.

4.3 Effect of Time Derivatives on Magnetic Field

The approximation of different derivates in case of different field values with finite differences in Maxwell's equations is explained in detail in Appendix 1. We calculate these terms before iteration to speed the process of simulation. The algorithm for calculation of H and e field values is generated

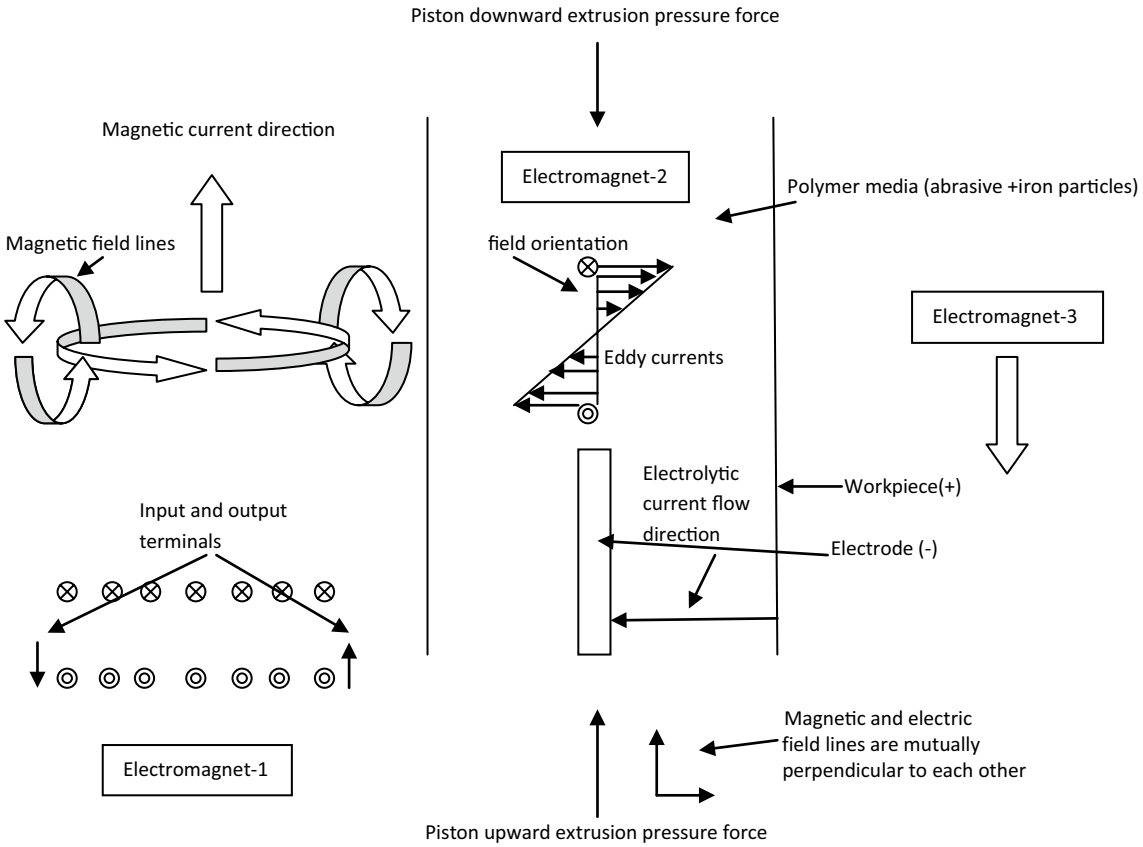


Fig. 13 Forces acting on abrasive particle inside the workpiece

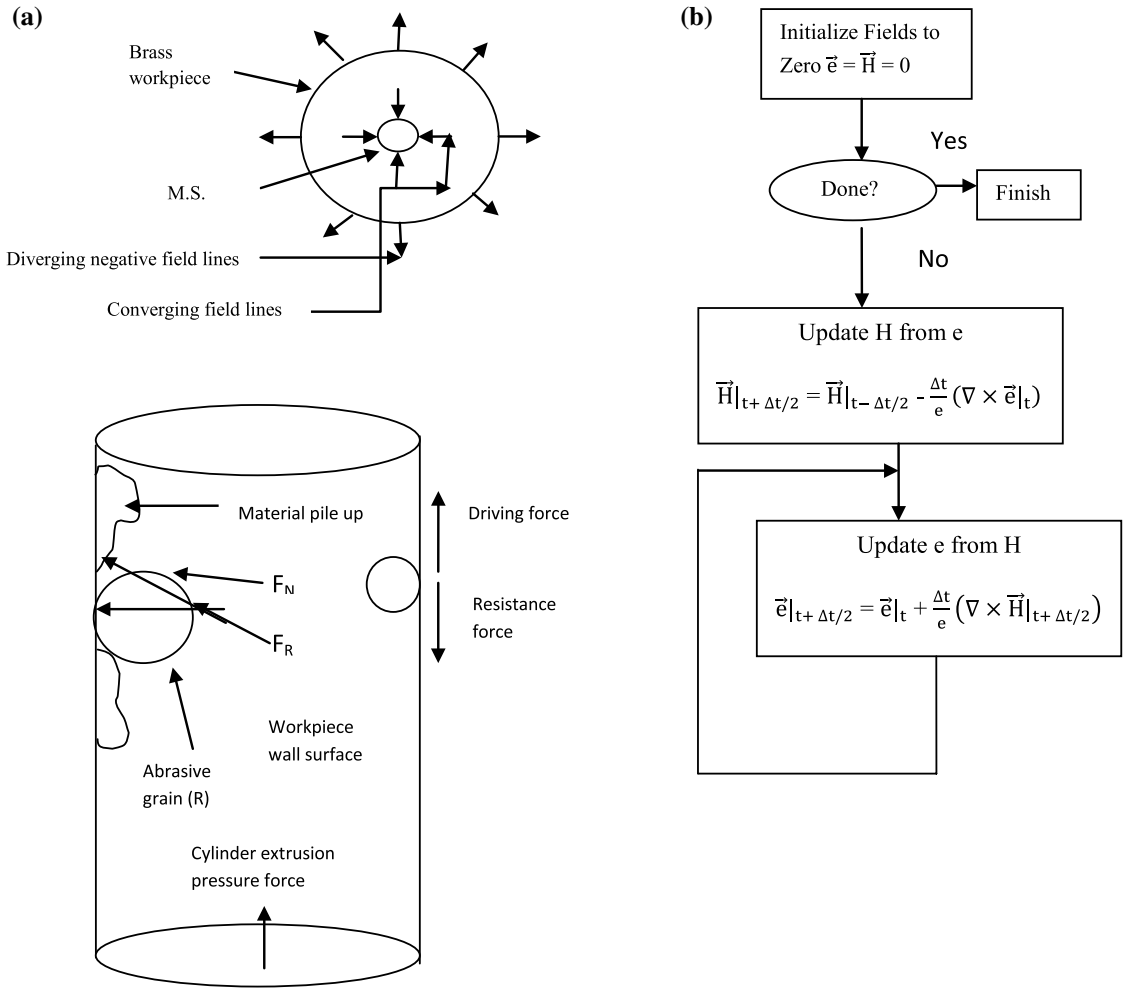


Fig. 14 **a** Different forces acting on the workpiece, **b** Time derivatives effect in magnetic field flow chart

which is shown as in Fig. 14b. The algorithm in which H and corresponding e values are calculated using iteration principle is generated. Initially $\vec{H}|_{t-\Delta t/2}=0$ and $\vec{e}|_t=0$. In the first iteration, $\vec{e}|_t \rightarrow \vec{H}|_{t-\Delta t/2}$. This suggests that the e field value at time t changes to H field value at half time step more than initial time value. Similarly the corresponding e and H values after every half time step value are explained. In the following iterations, $\vec{H}|_{t+\Delta t/2} \rightarrow \vec{e}|_t$ $\vec{H}|_{t+3\Delta t/2} \rightarrow \vec{e}|_{t+2\Delta t}$ and so on. The modeling included different laws i.e. Maxwell, Gauss, Faraday and Ampere law, etc. the above equations were utilized during modeling in Appendix 1. The electric field generated during the operation of magnetic field in hybrid AFM setup diverges from positive charges whereas convergence occurs in case of negative charges. Figure 14a shows the workpiece and tool electric field charge directions that are either converging or diverging in nature. If we stop the supply of current, there are no charges and hence loops are formed. The magnetic field produced during the working of prepared electromagnetic setup in hybrid AFM process always form closed loops. The magnetic field strength was calculated and was upto 1 Tesla.

This magnetic field causes the magnetic abrasive particles mixed in polymer media to attract towards the magnetized poles and hence hit the inner walls of brass workpiece, consequently improving the machining efficiency. The varying magnetic field with time is due to circulating electric field whereas the time varying magnetic fields induce circulating electric fields. The time derivative effect has been shown in Fig. 14b in the form of a flowchart.

4.4 Analysis of Flow of Media

When the abrasive laden polymer media is reciprocated between upper and lower cylinders, and hit the work surface, it exerts force and this force can be divided into different components like axial i.e. along the centerline axis, along the radius of circular job and also perpendicular to the radius i.e.

tangential components. Let k = 1, y = distance moved by media per stroke of machining = 650 mm, time/ stroke = 6 s.

hence, $\mu_o = 1 \times \left(\frac{650}{6}\right)^{0.8-1} = 0.3917$.

Let F_a, F_r and F_t are the axial, radial and tangential forces exerted by the polymer media on the abrasive particle, shown in Eq. (19). Extrusion pressure at which AFM machine is set = 30 bar = 30×10^5 Pa.

$F_a = p.A = 30 \times 10^5 \times 100.53/10^6 = 301.6$ N.

$F_r = p_m.A = 2 \times 10^5 \times 100.53/10^6 = 20.106$ N

$$F_t = F_a.\cos 45^o = 213.3 \text{ N} \tag{19}$$

where p_m is magnetic pressure = 2 Tesla = 2×10^5 Gauss. It can be seen that axial force is about 15 times the radial force, and 1.5 times the tangential force.

The cutting force, F_c is given in Eq. (20) and shear force is given by Eq. (21)

$$F_c = \sqrt{Fa^2 + Fr^2 + Ft^2}, \tag{20}$$

$$= \sqrt{301.6^2 + 20.1^2 + 213.3^2} = 369.42\text{N}$$

$$F_s = \tau_s.A \tag{21}$$

where F_s is the strength of workpiece and A is the projected area. For brass material, $\tau_s = 234$ MPa.

$F_s = 234 \times 100.5/10^6 = 0.0235$ N.

For positive condition of AFM action to occur, shear force must be lesser in magnitude as compared to cutting force. Since $F_c \geq F_s$, hence cutting occurs. The flow of semi-paste polymer media inside the workpiece is assumed as turbulent flow. The shear stress is given by Eqs. (22) and (23).

$$\tau = n\frac{du}{dy} + \mu\frac{du}{dy} \tag{22}$$

$$\tau = 0.9\frac{108.33}{650} + 48\frac{108.33}{650} = 8.05\,\text{Mpa}$$

$u^1 = v^1 = 43.2$ mm/s

$\tau = 1.1 \times 10^{-3} \times 43.2^2 = 2052.86$ MPa

$$\tau = \rho 1^2 \left(\frac{du}{dy}\right)^2 \tag{23}$$

$$= 1.1 \times 10^{-3} \times 260^2 \times 0.166^2 = 2.05 \text{ MPa}$$

$$\sqrt{\frac{\tau}{\rho}} = l\frac{du}{dy} \tag{24}$$

$\frac{du}{dy} = \sqrt{\frac{2.05}{1.1/1000}}/0.26 = 183.35\ s^{-1}$.

Shear velocity, $V_* = 0.4\ y\ \frac{du}{dy}$.

(23) = 19.06 mm/s

$2.5\ V_* \frac{dy}{y} = du$.

$2.5\ V_*\ \ln y = u + c$

Boundary conditions: at y = R, $u = u_{max}$

$$\frac{umax - u}{V*} = 5.75\ \log_{10}\frac{R}{y} \tag{25}$$

At y = R, $u_{max} = u = 43.2$ mm/s.

The above Eq. (25) is based on Prandtl velocity defect equation.

4.5 Computational Flow Analysis of Polymer Media in Hybrid Magnetic Assisted Abrasive Flow Machine Fixture

First of all, the nylon fixture for hybrid AFM process setup alongwith electromagnets was designed as shown in Fig. 15a, then modeling simulation and analysis of the flow of media was done. It included velocity, pressure, shear rate calculation, etc. There is constant density of media during operation. The density of abrasive laden polymer media is 740 kg/m^3. There is no effect on the density, it remains constant throughout the working operation. The shear rate variation with the flow passage is shown clearly in the Fig. 15b, alongwith scale on the left side ranging from 2638 to 12,006.74/s. As shown in the fixture, the shear rate value at entrance and exit lies between 1336.42/s to 2670.22/s, while all along the flow path the value was changing from 2638/s to 1336.42/s and the minimum shear rate value ranges from 2670.22/s to 4004.006/s which occurs at the junction of tapered portion of fixture and at the centre of workpiece location inside the fixture. This is due to the reason that the abrupt change in velocity occurs at this stage.

The pressure value ranges from − 297,880 Pa to 1.22 e + 07 Pa as shown in the scale in Fig. 15d. The maximum pressure value is at the centre and gradually decreases upto the exit cross section of the fixture and similarly this pressure change is seen from opposite side of the fixture, now the exit becomes the entrance of the fixture. The pressure value at the centre is 5277112 Pa.

The temperature variation is negligible during the media flow through the workpiece along the fixture path. It varies in the range 293.2 to 301.43 K. According to global coordinate system, the velocity of media variation is clearly depicted in the Fig. 15c, ranging from − 3.55 m/s to + 3.543 m/s, maximum being at the cross-section changes in the fixture and second most near the workpiece location. According to continuity equation of flow, as the cross section area increases, velocity of flow decreases and vice-versa, the minimum velocity being at the inlet and exit of the fixture.

Further, the top view of the machine fixture showing the inlet and outlet of the media flow, meshing and effect on

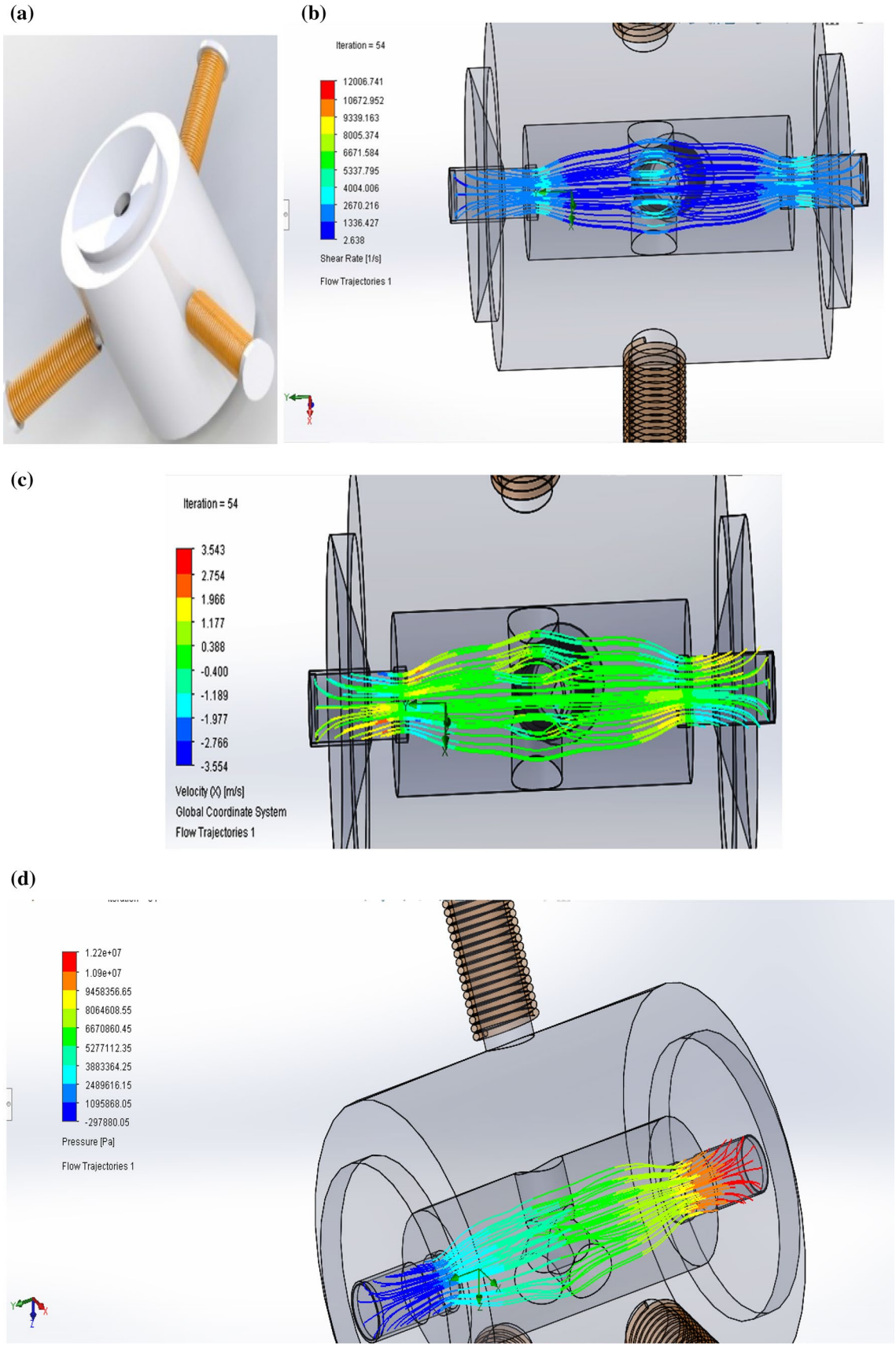


Fig. 15 **a** Design of AFM fixture, **b** shear rate, **c** velocity of media flow and (**d**) pressure variation

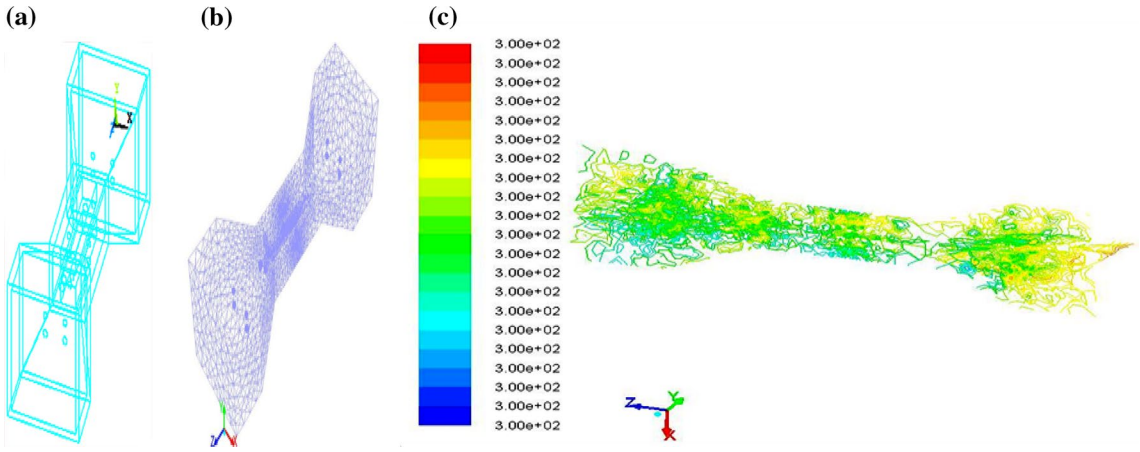


Fig. 16 **a** Top view of AFM, **b** Meshing of AFM, **c** Effect on temperature

temperature is shown in Fig. 16c. It can be clearly inferred that temperature is nearly constant when polymer media laden with abrasive particles pass through machine fixture. The value of temperature is given in Table 7. Before analyzing the effect of different parameters, i.e., temperature, pressure, shear rate, velocity, etc., the meshing of the flow of media between upper and lower extrusion cylinders is essential, as shown in Fig. 16b.

A small element is taken on the flow field and values of various factors such as location, velocity, temperature, pressure, shear stress and shear rate are given in the Table 7. It can easily be seen that temperature of media during its flow in the experimentation remains nearly constant, i.e., 293 K to 294 K. The velocity of flow of media varies from 0 to 2.008 m/s, this value obtained is nearly equal to that obtained from mathematical calculations. The shear rate changes from 6.155 s^{-1} to 2037s^{-1}, the drastic change may be due to the fact that firstly when media enters the fixture, the shear rate is low but when it comes in contact with workpiece surface, the shear rate value drastically increases due to the force exerted by solid surface onto the semi-solid of liquid layer of media. The pressure ranges from 11 to 67.5 Pa, which depends on machine parameter, i.e., extrusion pressure.

5 Conclusion

(1) The electrochemical and electromagnetic setup was employed successfully to make the traditional AFM setup hybrid in nature in order to improve the efficiency of the process. A machine setup alongwith electromagnet, media have been developed and characterized using FTIR and XRD analysis.

(2) The SBR media resulted in maximum material removal during experimentation, i.e., 3.88 mg when input parameters, i.e., electrolytic voltage, number of extrusion cycles and pressure were taken as 18 V, 4 and 10 bar respectively. The NR, NTR and SR media had intermediate effect of material removal but minimum removal of material was achieved in case of PBS media, i.e., 2.39 mg at 6 V voltage, 6 number of cycles and 30 bar pressure.

(3) The material removal was first increased with higher rod size but afterwards its increase was lesser. The surface plots obtained from RSM technique showed that MR obtained was 2.25 mg at 21 bar pressure and 7 number of cycles. As compared to conventional AFM setup, it was found that in ECMR-AFM setup, 34.5% and 17.8% improvement in % Ra and material removal, respectively, was obtained. It was found that MR was approximately 2.9 mg on an average when machining was done on traditional AFM process, while it increased upto 4.5 mg in prepared hybrid machine setup.

(4) Both mathematical and computational modeling has been done successfully. The magnetic flux and voltage calculations and electrochemical action has been explained alongwith time derivatives of magnetic field using Maxwell equations. The flux φ of magnetic field was found to be 300 G/A, magnetic field equal to 10 Gauss and angular momentum of iron particle due to rotation provided by motor was found to be 0.75 rad/s.

(5) Analysis of media flow both mathematically by calculating different forces that are shown by drawing a model and by using SolidWorks software, the flow was analysed using computational flow dynamics.

(6) The different optimization techniques i.e. Taguchi L9 OA, RSM, Minitab fuzzy logic and grey relational analysis were employed and implemented to validate the results obtained from the hybrid machine setup experimentation. The objective of maximum material

removal and minimum surface roughness was fulfilled consequently.

(7) The effect of different parameters viz. finishing cycles, different media, etc. on surface roughness improvement have been explained graphically in detail. It was found that 40% more reduction in surface roughness was achieved in 20% lesser number of cycles in hybrid ECMR-AFM setup.

Appendix 1: Laws Used in Magnetic Field and Time Derivative Approach

The time derivative approach in magnetic field generation is explained as under in Eqs. (26) and (27).

$$\nabla \times \vec{H}|_{t+\Delta t/2} = e_1 \frac{\vec{e}|_{t+\Delta t} - \vec{e}|_t}{\Delta t} \tag{26}$$

$$\vec{e}|_{t+\Delta t/2} = \vec{e}|_t + \frac{\Delta t}{e}\left(\nabla \times \vec{H}|_{t+\Delta t/2}\right) \tag{27}$$

where $\vec{e}|_{t+\Delta t/2}$ is field at next time step, $\vec{e}|_t$ is field at previous time step, $\frac{\Delta t}{e}$ is update coefficient and $\vec{H}|_{t+\Delta t/2}$ is curl of the other field at intermediate time step. The Maxwell Eq. (27) is used to determine electromagnetic forces and the direction of the magnetic lines. The equation can be denoted mathematically as:

$$\nabla.\mathrm{B} = 0 \tag{28}$$

where B is the magnetic flux density.

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{d\varphi}{dr}\right) + \frac{d^2\varphi}{dz^2} = 0 \tag{29}$$

where φ is the azimuthal angle, z is the height of the workpiece where the effect of magnetic field is produced. The differential Eq. (29) clearly explains the solution of azimuthal angle.

Gauss' Law:

$\nabla.\vec{S} = \theta$

$$\nabla.\vec{S} = \frac{\partial Sx}{\partial x} + \frac{\partial Sy}{\partial y} + \frac{\partial Sz}{\partial z} \tag{30}$$

$\nabla.\vec{B} = 0$

$$\nabla.\vec{B} = \frac{\partial Bx}{\partial x} + \frac{\partial By}{\partial y} + \frac{\partial Bz}{\partial z} \tag{31}$$

The Gauss' law is explained in Eqs. (30) and (31) Faraday's Law is explained in Eq. (32).

$$\nabla \times \vec{e} = -\frac{\partial \vec{B}}{\partial t}$$

$$\nabla \times \vec{e} = \left(\frac{\partial ex}{\partial y} - \frac{\partial ey}{\partial z}\right)\mathrm{bx} + \left(\frac{\partial ey}{\partial z} - \frac{\partial ez}{\partial x}\right)\mathrm{by} + \left(\frac{\partial ey}{\partial x} - \frac{\partial ez}{\partial y}\right)\mathrm{bz} \tag{32}$$

According to Ampere's law as shown in Eq. (33),

$$\nabla \times \vec{H} = \vec{i} + \frac{\partial \vec{S}}{\partial t} \tag{33}$$

where $\vec{i}$ is the current density. A B field induces an H field which is directly proportional to the permeability, as shown in Eq. (34).

$$\vec{B}(\mathrm{t}) = \left[\mu(\mathrm{t})\right].\vec{H}(\mathrm{t}) \tag{34}$$

where μ is the permeability index of the medium.

$$\vec{S}(\mathrm{t}) = \left[e_1(\mathrm{t})\right].\vec{e}(\mathrm{t}) \tag{35}$$

The above Eq. (35) denotes the induction of e field due to S field which is proportional to permittivity e_1.

References

1. Dixit, N., Sharma, V., & Kumar, P. (2021). Research trends in abrasive flow machining: A systematic review. *Journal of Manufacturing Processes, 64*, 1434–1461.
2. Fountas, N. A., & Vaxevanidis, N. M. (2021). Optimization of abrasive flow nano-finishing processes by adopting artificial viral intelligence. *Journal of Manufacturing Materials Process, 5*(1), 22–32.
3. Seyedi, S. S., Shabgard, M. R., Mousavi, B., & Heris, S. Z. (2021). The impact of SiC, Al2O3, and B2O3 abrasive particles and temperature on wear characteristics of 18Ni (300) maraging steel in abrasive flow machining (AFM). *International Journal of Hydrogen Energy*. https://doi.org/10.1016/j.ijhydene.2021.04.051
4. Tripathi, D. R., Vachhani, K. H., Bandhu, D., Kumari, S., Kumar, V. K., & Abhishek, Kumar. (2021). Experimental investigation and optimization of abrasive waterjet machining parameters for GFRP composites using metaphor-less algorithms. *Materials and Manufacturing Processes*. https://doi.org/10.1080/10426914.2020.1866193
5. Routara B, Nanda B.K, Sahoo A.K, Thatoi D.N. Optimisation of multiple performance characteristics in abrasive jet machining using grey relational analysis. International Journal of Manufacturing Technology and Management 2021; 24(1/2/3/4): 4 – 22
6. Kumar, R., Modi, A., Panda, A., Sahoo, A. K., Deep, A., Behra, P. K., & Tiwari, R. (2019). Hard turning on jis s45c structural steel: An experimental, modelling and optimisation approach. *International Journal of automotive and mechanical engineering, 16*(4), 23–33.
7. Mohanty, S., Routara, B. C., Nanda, B. K., Das, D. K., & Sahoo, A. K. (2018). Study of machining characteristics of Al-SiCp12% composite in nano powder mixed dielectric electrical discharge machining using RSM. *Materials Today Proceedings, 5*(11), 22581–25590.

8. Routara, B. C., Das, D., Satpathy, M. P., Nanda, B. K., Sahoo, A. K., & Singh, S. S. (2020). Investigation on machining characteristics of T6-Al7075 during EDM with Cu tool in steady and rotary mode. *Materials Today Proceedings, 26*(2), 2143–2150.
9. Banik D, Rahul, Kar G, Debnath B, Routara B.C, Sahoo A.K, Kochar D. (2019) Machining Performance Optimization During Electro Discharge Machining on Titanium (Grade 4): Application of Satisfaction Function and Distance-Based. Advances in Industrial and Production Engineering 2021: 535–542
10. Thamizhvalavan, Y. N., & Arivazhagan, S. (2021). Abrasive water jet machining of al6063/b4c/zrsio4 hybrid composites: A study of machinability and surface characterization analysis. *Silicon*. https://doi.org/10.1007/s12633-020-00888-2
11. Baraiya, R., Babbar, A., Jain, V., & Gupta, D. (2020). In-situ simultaneous surface finishing using abrasive flow machining via novel fixture. *Journal of manufacturing processes, 50*, 266–278.
12. Han, S., Salvatore, F., Rech, J., & Bajolet, J. (2020). Abrasive flow machining (AFM) finishing of conformal cooling channels created by selective laser melting (SLM). *Precision Engineering, 64*, 20–23.
13. Wei, H., Peng, C., Gao, H., Ping, X., & Wang, X. (2019). On establishment and validation of a new predictive model for material removal in abrasive flow machining. *International Journal of Machine Tools and Manufacture, 138*, 66–79.
14. Wei, H., Peng, C., Gao, H., Wang, X., & Wang, X. (2019). On establishment and validation of a new predictive model for material removal in abrasive flow machining. *International Journal of Machine Tools & Manufacture, 138*, 66–79.
15. Wei, H., Gao, H., & Wang, X. (2019). Development of novel guar gum hydrogel based media for abrasive flow machining shear thickening behavior and finishing performance. *International Journal of Mechanical Sciences, 157–158*, 758–772.
16. Singh, P., Singh, L., & Singh, S. (2020). Manufacturing and performance analysis of mechanically alloyed magnetic abrasives for magneto abrasive flow finishing. *Journal of manufacturing processes, 50*, 161–169.
17. Zhao, J., Huang, J., Wang, R., Peng, H. R., Hang, W., & Ji, S. (2020). Investigation of the optimal paramaters for the surface finish of K9 optical glass using a soft abrasive rotary flow polishing process. *Journal of manufacturing processes, 49*, 26–34.
18. Zhang, J., Chaudhari, A., & Wang, H. (2019). Surface quality and material removal in magnetic abrasive finishing of selective laser melted 316 L stainless steel. *Journal of manufacturing processes, 45*, 710–719.
19. Ali, P., Dhull, S., Walia, R. S., Murtaza, Q., & Tyagi, M. (2017). Hybrid abrasive flow machining for nano-finishing- a review. *Materials Today Proceedings, 4*(8), 7208–7218.
20. Dhull, S., Walia, R. S., Murtaza, Q., & Niranjan, M. S. (2019). Electrochemo- magneto abrasive flow machining setup fabrication and experimental investigation of the process alongwith mathematical modelling and simulation. *Independent Journal of Management & Production, 10*(8), 1834–1849.
21. Brar BS, Walia RS, Singh VP, Sharma M. Helical abrasive flow machining (hlx-afm) process. International Journal of Surface engineering and Materials Technology 2012; 604–608.
22. Ravi Sankar, M., Mondal, S., Ramkumar, J., & Jain, V. K. (2012). Experimental investigations and modelling of drill bit guided abrasive flow finishing (DBG-AFF) process. *International Journal of Advanced Manufacturing Technology, 42*, 678–688.
23. Sharma AK, Kumar P and Rajesha S. An improved ultrasonic abrasive flow machining. Patent number 3578/DEL/201, India.
24. Singh, S., & Shan, H. S. (2002). Development of magneto abrasive flow machining process. *International Journal of Machine Tool & Manufacture, 42*, 953–959.
25. Charles C. Family Firm to Modern Multinational - Norton Company, a New England Entreprise, Cambridge, MA, Harvard University Press.
26. Petri, K. L., Billuo, R. E., & Bidando, B. (1998). A neural network process model for abrasive flow machining operation. *Journal of manufacturing systems, 17*(1), 52–64.
27. Silva, A. K. M. D., & Geough, J. A. M. (2000). Computer applications in unconventional machining. *Journal of Materials Processing Technology, 107*, 276–282.
28. Kumar, P., Tailor, B., Agrawal, A., & Joshi, S. S. (2013). Evolution of electrochemical finishing processes through cross innovations and modeling. *International Journal of Machine Tools & Manufacture, 66*, 15–36.
29. Paul, L., Shekhar, S., & Hiremath, S. (2013). Response surface modelling of micro holes in electrochemical discharge machining process. *International conference on design and manufacturing, 64*, 1395–1404.
30. Zaborski, S., Łupak, M., & Poro, D. (2004). Wear of cathode in abrasive electrochemical grinding of hardly machined materials. *Journal of Materials Processing Technology, 149*, 414–418.
31. Jain, R. K., Jain, V. K., & Kalra, P. K. (1999). Modelling of abrasive flow machining process: A neural network approach. *Wear, 231*, 242–248.
32. Venkatesh, G., Sharma, A. K., & Singh, N. (2015). Simulation of media behaviour in vibration assisted abrasive flow machining. *Simulation Modelling Practice and Theory, 51*, 1–13.
33. Sankara M, Babu AG. Effect of reinforcement particles on the abrasive assisted electrochemical machining of Aluminium-Boron carbide-Graphite composite. 12th Global congress on manufacturing and management. Procedia Engineering 2006; 97: 381 – 389.
34. Dabrowski, L., Marciniak, M., Wieczarek, W., & Zygmunt, A. (2006). Advancement of abrasive flow machining using an anodic solution. *Journal of New Materials for Electrochemical Systems, 9*, 439–445.
35. Singh, R., Walia, R. S., & Suri, N. M. (2015). Parametric optimization of Centrifugal- magnetic force assisted abrasive flow machining process using utility concept. *International Journal of Research in Engineering and Technology, 4*(8), 2321–2230.
36. Jui, S. K., Kamaraj, A. B., & Sundaram, M. M. (2015). High aspect ratio micromachining of glass by electrochemical discharge machining. *Journal of Manufacturing Processes, 15*, 460–466.
37. Uhlmann, E., Schmiedel, C., & Wendler, J. (2015). CFD simulation of the abrasive flow machining process. *Procedia CIRP, 31*, 209–214.
38. Di Corleto J. Innovations in Abrasive Products for Precision Grinding. Precision Grinding and Finishing Conference, Gorham 2001.
39. Jain, R. K., & Jain, V. K. (1999). Simulation of surface generated in abrasive flow machining process. *Robotics and computer integrated manufacturing, 15*, 403–412.
40. Yan, B. H., Wang, C. C., & Lin, Y. C. (2006). Feasibility study of rotary electrical discharge machining with ball burnishing for Al2O3/6061Al composite. *International Journal of Machine Tools & Manufacture, 40*, 1403–1421.
41. Tzeng, H. Z., Yan, B. H., Hsu, R. T., & Chow, H. M. (2007). Finishing effect of abrasive flow machining on micro slit fabricated by wire-EDM. *International Journal of Advanced Manufacturing Technology, 34*, 649–656.
42. Zhang, M., & Jin, D. G. (2009). EDM performance of electroformed Cu-ZrB2 shell electrodes. *Rapid Prototyping Journal, 15*(2), 150–156.
43. Chryssolouris G., Guillot M. Modelling of machining process using neural network. Trans. ASME 1990 J. Eng. Ind, Vol. 112, pp. 122–131.

44. Ghoshal, B., & Bhattacharyya, B. (2015). Investigation on profile of microchannel generated by electrochemical micromachining. *Journal of Materials Processing Technology, 222*, 410–421.
45. Jahan, M. P., Rahman, M., & Wong, Y. S. (2011). A review on the conventional and micro-electrodischarge machining of tungsten carbide. *International Journal of Machine Tools & Manufacture, 51*, 837–858.
46. Jiang, B. Y., Lan, S. H., & Ni, J. (2009). Modeling and experimental investigation of gas film in micro-electrochemical discharge machining process. *International Journal of Machine Tools & Manufacture, 90*, 8–15.
47. Yang, C. T., Song, S. L., Yan, B. H., & Huang, F. Y. (2006). Improving machining performance of wire electrochemical discharge machining by adding SiC abrasive to electrolyte. *International Journal of Machine Tools & Manufacture, 46*, 2044–2050.
48. Lin, Y. C., & Lee, H. S. (2009). Optimization of machining parameters using magnetic-force-assisted EDM based on gray relational analysis. *International Journal of Advanced Manufacturing Technology, 42*, 1052–1064.
49. Niranjana MS, Jha S. Flow Behaviour Of Bidisperse MR Polishing Fluid And Ball End MR Finishing. 3rd International Conference on Materials Processing and Characterisation Procedia Materials Science 2009; 6: 798 – 804.
50. Das, M., Jain, V. K., & Ghoshdastidar, P. S. (2008). Analysis of magnetorheological abrasive flow finishing (MRAFF) process. *International Journal of Advanced Manufacturing Technology, 38*, 613–621.
51. Singh, A. K., Jha, S., & Pandey, P. M. (2012). Nanofinishing of a typical 3D ferromagnetic workpiece using ball end magnetorheological finishing process. *International Journal of Machine Tools & Manufacture, 63*, 21–31.
52. Kala, P., & Pandey, P. M. (2015). Comparison of finishing characteristics of two paramagnetic materials using double disc magnetic abrasive finishing. *Journal of Manufacturing Processes, 17*, 63–77.
53. Vaishya R, Walia RS, Kalra P. Design and Development of hybrid electrochemical and centrifugal force assisted abrasive flow machining. 4th International Conference on Materials Processing and Characterization, Materials Today Proceedings 2015; 2: 3327 – 3341.
54. Roades LJ. Abrasive flow machining. Technical Paper of Society of Manufacturing Engineers (SME): 89–145 .
55. Kurita, T., & Hattori, M. (2006). A study of EDM and ECM/ECM-lapping complex machining technology. *International Journal of Machine Tools & Manufacture, 46*, 1804–1810.
56. Pal, P., & Jain, K. K. (2018). Computational simulation of Abrasive Flow Machining for two dimensional models. *Materials Today Proceedings, 5*(2), 12969–12983.
57. Gupta, G. A., Ansari, I. A., Ramkumar, J., & Kar, K. K. (2021). Rheological characterization of newly developed fly-ash mixed polymer media and its finishing performance through abrasive flow machining. *Cleaner Engineering and Technology, 2*, 85–100.
58. Fu, Y., Gao, H., Yan, Q. S., & Wang, X. P. (2019). A new predictive methodof the finished surface profile in abrasive flow machining process. *Precision Engineering, 60*, 497–505.

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



Sachin Dhull has done Ph.D. in advanced manufacturing process from Delhi Technological University, Delhi and is currently working as an assistant professor in Maharaja Surajmal Institute of Technology, Delhi, India.



R. S. Walia is Professor & Head, Production and Industrial Engineering, Punjab Engineering College, Chandigarh, India. He has done Ph.D. from IIT Roorkee. His research interests include Advanced Manufacturing Processes, Metal Matrix Composites, Thermal Coating, Surface Engineering, Optimization, Additive Manufacturing, Industrial Engineering, Supply Chain Management, Ergonomics.



Qasim Murtza is Professor in mechanical engineering department, Delhi Technological University, Delhi. His PhD specialization is in Production Engg, Materials, Metal Coating -Advance machining.



M. S. Niranjan is Associate Professor in mechanical engineering department, Delhi Technological University, Delhi. His qualification includes: B.Tech (IIT-Roorkee), M.Tech (IIT-Roorkee), PhD (IIT Delhi). His specialization is in Production & Industrial Engg.

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

Extracting Notes From Youtube Video

Manoj Kumar
Department Of Computer Engineering
Delhi Technological University
New Delhi, India
mkumarg@dce.ac.in

Sumit Gaurav
Department of Computer Engineering
Delhi Technological University
New Delhi, India
sumitgaurav_2k17co349@dtu.ac.in

Suhail Akhtar
Department of Computer Engineering
Delhi Technological University
New Delhi, India
suhailakhtar_2k17co345@dtu.ac.in

Siddharth Varshney
Department of Computer Engineering
Delhi Technological University
New Delhi, India
siddharthvarshney_2k17co340@dtu.ac.in

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498392

***Abstract*—In the world of technology, every B.tech student in our college uses Youtube to study the night before the exam, it will be of better use if there is an algorithm that can also provide notes for that video to look for last-minute revision. Hand-written notes are to be carried in a notebook, tedious to maintain for a long time. If a student wishes to pause and take notes, it will consume double the regular time of that streaming video. Those notes might or might not be used in the future or maybe not give proper context if the original content of the video is forgotten. In this project we are extracting frames from a presentation-based youtube video after every fixed time interval, using optical character recognition techniques on these images to make notes of the content written in a youtube video. Users will give the link of the desired youtube video and a pdf/doc file containing all the sentences shown in a video will be in the output. Students will just have to edit and delete some extra material that is not needed in notes. In the upcoming era education will bend more towards paperless, so to maintain the standards all the notes will be generated using this concept. Though it may sound simple it is a bit complex as to know when to extract data from given frame or how to handle cases when some person come in between the video. All this challenges will be major challenges in the future and will require a solution our idea is the first step towards this future .**

***Index Terms*—YouTube, Key-Frame extraction, Optical Character Recognition, String matching, Image Structural Similarity**

I. Introduction

This is the process of creating a document based on the texts shown in the video. This technology can be used in a huge variety of segments, such as making notes for regular classes, instead of wasting time on writing down calculations of mathematical equations, just formulas can be remembered and the rest can be left for this program to write down all the text shown on the screen written by a teacher. Video-based Notes extraction or note creation is a new concept that is not performed by any known software in the author's knowledge.

The motivation for this project came from studying the night before an exam and wasting time copying notes from the study video for future use. Most of the content shown in a video is needed only to remember the context of a topic or to remember some keywords for an answer. Making these types of notes takes a lot of time. Automating this whole notes making process just providing a link to the video will give a boost to the amount of content that could be studied for a limited amount of time.

Several steps have been performed in creating the program. The first step is extracting frames from a video that would be most refined or we say frames that contain the most useful content to be written in notes. We have calculated the Frame Rate per second.

This method is known as Key-Frame extraction. This method has been used in advertisement industries[1] for many years such as in Google and Amazon advertisements. The key-Frame extraction method studies consecutive frames of a video and analyses them against the previously selected frame of the video stored in the database. As soon as it calculates a particular frame is distinct enough from the last one, it stores that frame for further processing. The key-Frame method generally is applied in videos that have high feature differences, the challenge we are facing in applying this algorithm in our work is that the videos are presentation based in which theme is the same throughout the video and only the basic content of the video i.e. the English sentences are changing without a single diagram, throughout the video. In these types of videos, there is not much of a feature difference so after the first selected frame, the key-frame extraction method will not select any frame for further processing. Thus, we are extracting frames at a regular time interval.

The second step of the project is using Optical Character Reader technology. Optical Character Reader is also known as OCR in short is the conversion of graphical characters present in a sentence or in any form in an image such as scanned image, clicked the image, frame captured from a video to digital format of the same characters, and sentences. The need for the technique arose as to when archaeologist

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:56:55 UTC from IEEE Xplore. Restrictions apply.

were converting old manuscripts in digital format, it took a huge amount of time for such a task manually, there was a need to automate this task to save time. In our project, we have used a google python library named Tesseract as an OCR. Tesseract is free, modifiable, and gives us fairly good results on simple English text images. There are some defects this open-source library can face, such as it gives poor results in the colored background and stilted sentences. These defects are removed by binarization and image rotation during the image preprocessing phase before OCR. We have excluded these techniques.

This OCR-converted digital text is not totally unique in content between multiple frames. Redundancy has to be removed from the digital text that is going to be printed in the notes. The third step will contain an algorithm that is created which compares text from consecutive previous frames and removes the common data present in the previous frame from the current frame text.

II. METHODOLOGY

A. FRAME EXTRACTION

In this procedure, we extracted frames from different videos and compared each extracted frame to the consecutive extracted frame, and obtained a percentage of similarity among them which is qualified as a similarity index. We used open libraries of python such as "skimage.metrics.structural_similarity" and "cv2.ORB_create()", these libraries compared images based on their features and gave the result as how much percentage of the two images are exactly equal.

ORB Algorithm:- ORB stands for Oriented FAST and BRIEF. Fusion of FAST and BRIEF descriptor along with some modification mainly optimization led to the formation of ORB. FAST is a Feature from the Accelerated segment test which is used to detect features from provided images. Fast doesn't compute orientation and descriptors so for this BRIEF(Binary Robust Independent Elementary Features) is used. ORB rotate BRIEF according to the key points. ORB is the best alternative for SURF(Speed Up Robust Features) and SIFT(Scale Invariant Feature Transform). ORB has better performance in feature extraction, computation cost, and matching performance as compared to the former two.[14]

Structural_Similarity:- Structural_ similarity often written as ssim is used to calculate the Structural Similarity Index. Structural Similarity Index is a numeric value calculated between two images. Its value ranges from -1 to +1. A value of +1 indicates that both the images are similar with little to no difference and a value of -1 indicates that both the images are different. For most of the time, the range is normalized from 0 to 1 where the end carries the same meaning. Before the structural similarity index, the mean squared error was used to calculate the difference between two images because it is easy to implement but the accuracy is not up to the mark, that's why the structural similarity index is used nowadays.[15]

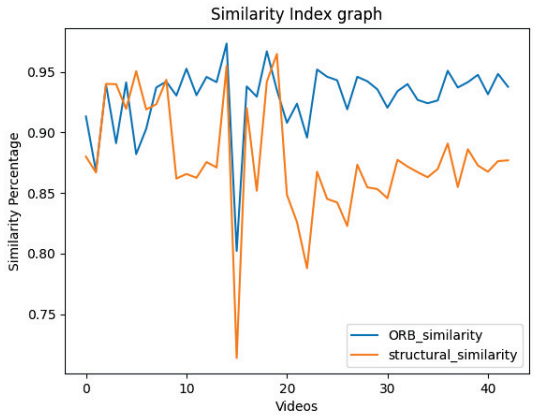


Fig. 1. Similarity Index Graph

About 45 videos were analyzed in this procedure. Videos were divided into two parts, which contain graphical images and another which contained only text. The average similarity index of frames which contain only text was very high as compared to the set of videos that contained images in some parts. We are taking an average of all the similarity indices of all comparisons between frames of a single video, then taking an average of all the indices computed for each video in a set. The Average Similarity Index of non-graphical videos was approx 87% and the set of videos with graphical content was approx 44%. This difference occurs as frames that contain images differ highly from frames with no image which gives the two frames a very low similarity index value thereby decreasing the overall similarity in different frames of that video and giving a false indication that the Key-frame extraction method can work, therefore we excluded graphical videos for computing result in this part.

By this result, we came to conclude that the Key-Frame Extraction method would not be a good option to extract the computable frames. Thus we are extracting frames at a constant time of 5 sec. This time difference to extract consecutive frames is based solely on a random value, it can be changed as per the usage of the system.

TABLE I
FRAME SIMILARITY TABLE

Algorithms / video_types	SSIM	ORB_Sim
No_Image_videos	0.92	0.87
Image_Videos	0.52	0.44

B. Optical Character Reader

Tesseract is the library used for reading characters from image and converting into digital format.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:56:55 UTC from IEEE Xplore. Restrictions apply.

Tesseract OCR:-
Tesseract is an open source OCR engine. In recent times its popularity among developers has increased significantly. It was originally developed by HP in the 1980s, and was finally made open source in 2006. Its performance is far better than similar libraries because it supports more than 100 languages and can be trained to learn new languages with higher accuracy. Tesseract is used for reading text from binarized images. It is used for text detection in mobile devices, in video as we are using and in detecting spam mails in gmail.In python this library is present with the name pytesseract. A simple example can be given with the following image.



Fig. 2. Figure shows how baseline are created.

As you can see it has extracted the text in red. Some of the methods used by tesseract library to detect texts are discussed below:-

Baseline fitting: It detects the lines which are straightly written by partitioning the line with the reasonably equidistant line for the original straight line.[13]

Fixed pitch detection: This is another method which is used by the tesseract library. In this it scans the entire line to check whether they are fixed pitch or not. Whenever it detects the fixed pitch texts, it slices the words into individual characters with the help of pitch. Following which it then disables other methods on it like associator and then chopper which helps it in finding given words correctly. Following figure will illustrate it perfectly.[13]



There are many more methods which are used by tesseract but these are the early ones and form the foundation of tesseract library.

C. Identifying New Sentences Between Two Slides

When we see any lecture video online we often see text on board changes gradually that is we don't see a 100% switch in text and so if we just copy everything from slide to notes it will contain a very high amount of duplicate text so in order to obtain new text from newer slide we require an algorithm to do so.
So to understand our algorithm let us define few terms :-

String1 :- It represents the text we got from the last frame.
String2 :- It represents the text we got from the current frame.
match_points :- match point is given to a pair of strings and high match_point represents a pair of strings that are very similar and low match_point shows less matching strings.
Method used for calculating match_points :-
To find match_points we are taking three different type of parameters :-
Uni :- It represents how many times a character has appeared in a string .
Example
string:-"ssumit"
Uni stores:- s-2,u-1,m-1,i-1 and t-1

Di :- It represents how many times a pair of characters has appeared in a string.
Example
string:-"siddd"
Di stores:- si-1,id-1 and dd-2

Tri :- It represents how many times a pair of three characters has appeared in the string.
Example
string:-"suhauha"
Tri stores:- suh-1, uha-2, hau-1 and auh-1

Match_points = number of uni matched between two strings + 2* number of di matched between two strings + 3* number of Tri matched between two strings.

Now we will take a different starting position from string1 and compare the first 20 letters from that position to the first 20 letters of string2 after finding all the possible values of match_points .

Now if we have a highest match_point value greater than equal to 30 we will say that yes there is some overlap otherwise we will consider it as 0 overlap and copy the entire string2 to our notes. Now in case of overlap the highest match_point value giving position is considered as the point from where overlap is starting .

Now from that position we will keep taking a block of 10 size and match it with corresponding block in string2 and if at any point mismatch in two blocks is greater than or equal to thirty percentage we will break our process at that point and copy entire string2 from that point and if this never takes place we will copy text from that position where we are left in string2.

Below is the example of processing of 2 frames and it's output:

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:56:55 UTC from IEEE Xplore. Restrictions apply.

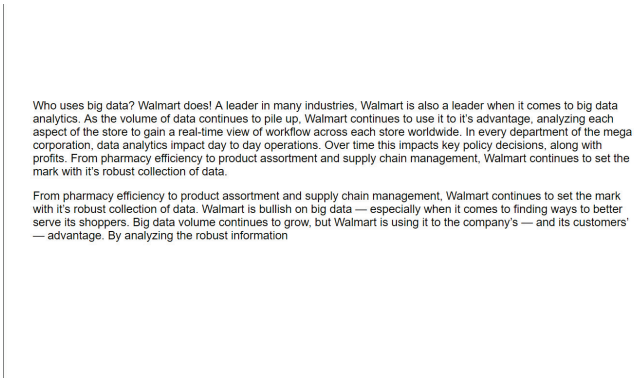


Fig. 3. Frame 1

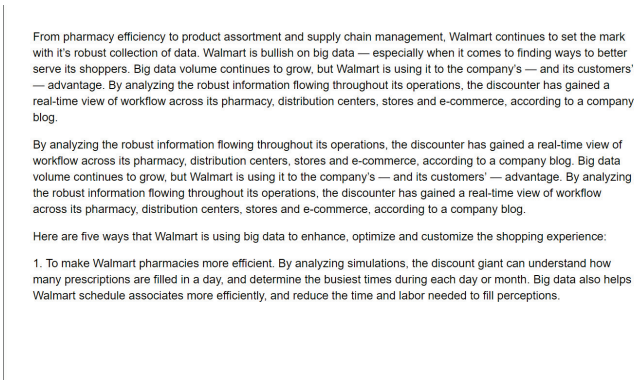


Fig. 4. Frame 2

Above consecutive frames give the following output:

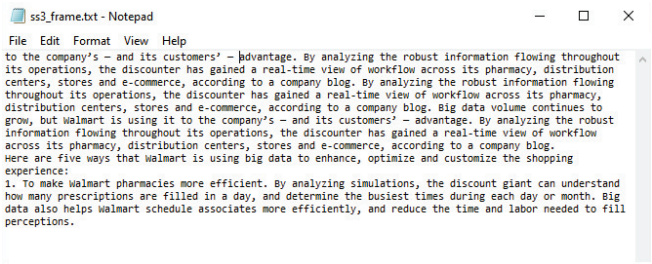


Fig. 5. Frame 3

In Fig.5 the output of the above frames comes with 7.3% repetition and 92.7% original text in the second frame. All the text which was present in frame 2 from frame 1 has been omitted, thus the final text file which will be generated will have maximum unique and minimum redundancy needed by the user.

III. RESULTS

A. *Comparing our output with original text:-*

The validity of our algorithm can be easily verified by comparing the length of our output with original output length . Since we are trying to add text from output generated by OCR the only concern left for us is , if we are able to copy only unique texts from generated output . To do so we are using a very simple algorithm that is to check by what percentage we are writing extra words or by what percentage we are writing less words. Though this may sound simple but it is effective and easy way to check the accuracy.

Algorithm:-

Percentage of mismatch = abs(length of generated output/length of original text - 1)*100

Example :-

Original text :- My name is Sumit Gaurav

Generated text :- My name is Sumit mit Gaurav

% mismatch = |27/23 - 1 |*100 = 17.39

Although this percentage seems so high in this sample but for larger texts this % will drop significantly as the words reused remain more or less the same but new words added generally increase significantly.

Example:-

Original text length :- 1200

Generated text length : - 1250

%mismatch = |1250/1200 -1|*100 = 4.16%

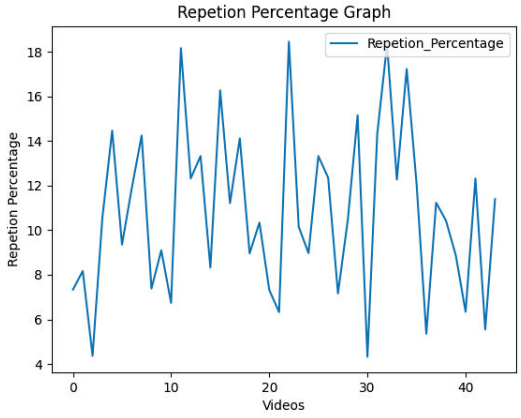


Fig. 6. Frame 1

For 45 videos, original text files and converted text files were compared according to the method above, giving us an average of about 10.8% repetition in text which frames originally had much more repetition.

Repetition varies from 4% to 19% across 45 videos. This means that all the necessary data is present in the text file present in the video with a little more due to repetition from consecutive frames.

IV. CONCLUSION

We started by extracting key-frames which contain most of the data related to the text part of video, but, due to high similarity in consecutive frames, we drop the idea for future

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:56:55 UTC from IEEE Xplore. Restrictions apply.

scope.We extracted frames every 5 seconds. OCR was applied to the given frames to convert text from digital form into string format for further processing. The text was compared among consecutive frames to remove redundant text which is common in both frames so that maximum unique content could be delivered into the final text file for the user. Applying the given model on the videos we were able to successfully extract text with 10% overlap which can be removed by the user manually. Videos containing graphical contents can not be processed due to OCR and phase III equation limitations.

V. FUTURE SCOPE

Key-Frame extraction method could be improved and will work with videos containing images to extract frames which are most data. With better OCR technology we can also include videos which contain images and videos in which a person is teaching on board with repetition aim of less than 5%. In coming era as we are relying more on technology, students will less likely make handwritten notes. Note making will done through artificial intelligence with more advanced feature.

Future scope can be broadly divided into three categories:
1. At the moment our project does not handle presentation with human body involvement that is if a human appears in front of presentation it will be a huge issue for us as we will not be able to get proper text but in future we would like to change it such that we can choose frames such that it involves minimum presence of human body and hence improving quality of notes.

2. There can be cases where instead of teaching through presentations, blackboard is being used. In this case our priority is to identify the handwriting of the teacher and then convert each note provided by the teacher into separate notes.

3. Along with the frame extraction our project can be expanded to also convert the speech of the instructor into text. For example consider that the teacher takes a pause and is giving a speech now that speech can be converted into text. From converted speech only the optimal part will be taken and can be merged with the optimal notes generated from the frames of the video. This will give the complete note of the entire lecture.

REFERENCES

[1] Karez Hamad,Mehmet Kaya."A Detailed Analysis of Optical Character Recognition Technology". International Journal of Applied Mathematics Electronics and Computers 4(Special Issue-1):244-244.

[2] Sheena, N.K.Narayanan. "Key-frame Extraction by Analysis of Histograms of Video Frames Using Statistical Methods."

[3] Yunyu Shi, Haisheng Yang, Ming Gong, Xiang Liu, and Yongxiang Xia."A Fast and Robust Key Frame Extraction Method for Video Copyright Protection."

[4] Gianluigi Ciocca Raimondo Schettini. " Dynamic key-frame extraction for video summarization". The International Society for Optical Engineering.

[5] "KEY FRAME EXTRACTION METHODS" By Israa Hadi Ali and Talib T Al-Fatlawi

[6] Z. Lu and Y. Shi, —Fast video shot boundary detection based on SVD and pattern matching,‖ IEEE Trans. Image Processing, Vol. 22, No. 12, 2013, pp. 5136-5145.

[7] Rachida Hannane, Abdessamad Elboushaki, Karim Abdel 1, P. Naghabhushan and Mohammed Javed, —An efficient method for video shot boundary detection and keyframe extraction using SIFT-point distribution histogram‖, International Journal of Multimedia Information Retrieval, Vol. 5, No. 2, 2016,pp. 89-104

[8] W. Hu, N. Xie, L. Li, X. Zeng, and S. Maybank, —A survey on visual content-based video indexing and retrieval‖, IEEE Transactions on Systems, Man, and Cybernetics—part c: Applications and Reviews, Vol. 41, No. 6, 2011, pp. 797 819.

[9] T. He, W. Huang, Y. Qiao, and J. Yao, "Text-attentional convolutional neural network for scene text detection," IEEE Transactions on Image Processing, vol. 25, no. 6, pp. 2529–2541, 2016.

[10] M. Jaderberg, K. Simonyan, A. Vedaldi, and A. Zisserman, "Reading text in the wild with convolutional neural networks," International Journal of Computer Vision.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in CVPR, 2016.

[12] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in Advances in neural information processing systems.

[13] An Overview of the Tesseract OCR Engine. https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/33418.pdf

[14] https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_feature2d/py_orb/py_orb.html

[15] Structural Similarity https://en.wikipedia.org/wiki/Structural_similarity

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:56:55 UTC from IEEE Xplore. Restrictions apply.

ACADEMIA
Accelerating the world's research.

IRJET-MFeasibility Study on Production of Hydrogen from Renewable Sources

IRJET Journal

IRJET

Cite this paper

Downloaded from Academia.edu

Get the citation in MLA, APA, or Chicago styles

Related papers

Download a PDF Pack of the best related papers



IRJET- Pyrolysis of Biomass and the use of Pyrolysis Oil as an Alternate Fuel: A Review
IRJET Journal

Production of Hydrogene and Environmental Impact of Storage the Energy Produced by Renewable En...
International Journal of Scientific Research in Science and Technology IJSRST

A comparative overview of hydrogen production processes
Pavlos Nikolaidis

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 06 | June 2021 www.irjet.net p-ISSN: 2395-0072

Feasibility Study on Production of Hydrogen from Renewable Sources

Saurav Yadav[1], Saurabh Agase[2], Dr. J.P. Kesari[3]

[1,2]Student, Dept. of Mechanical Engineering, Delhi Technological University, New Delhi, India
[3]Associate Professor, Dept. of Mechanical Engineering, Delhi Technological Engineering, New Delhi, India

---

Abstract - *Hydrogen is considered to be a clean fuel. It can be obtained from various domestic resources, such as biomass, natural gas, nuclear power, and renewable energies like solar and wind. Due to the growing need for hydrogen, the hydrogen economy is rising on the political agenda. The natural occurrence of hydrogen cannot satisfy the present need. Hence there is a wide gap between hydrogen present on earth and the current hydrogen requirement. Hydrogen energy is the zero carbon emission energy at the end but, it depends on the production pathway and the energy used for its production. Commercial Hydrogen is produced in industries through various methods. Most hydrogen in India is produced through reforming methane, resulting in significant carbon dioxide emissions. This paper examines the current world hydrogen energy scenario, the methods of production of hydrogen using non-renewable sources, and the economic aspect of hydrogen production through various ways like biomass gasification, Pyrolysis, fermentation, bio photolysis, and electrolysis process.*

Key Words: Hydrogen, Renewable Sources, Biomass Gasification, Pyrolysis, Fermentation, Bio Photolysis.

1. INTRODUCTION

Hydrogen is the most widespread chemical element on the earth and is molecular di-hydrogen (H2). It is expected to be an important fuel in the future that will positively affect the quality of atmospheric air. It can be deployed as an energy carrier to store, move, and deliver energy produced from other sources. It can be obtained from several sources, both renewable and non-renewable, by various processes. Today's most common methods are natural gas reforming (a thermal process), electrolysis, and other methods include other thermal methods and biological processes.

Developing energy systems for the future will be based on emission-free and low carbon sustainable energy sources. Hydrogen as a fuel will play a significant role in solving climate change and energy usage problems.

1.1 World Hydrogen Energy Scenario

Global hydrogen production and consumption is over 55 million tonnes per year [1]. Hydrogen technologies intensified in 2019, is awakening keen interest among policymakers. The fuel cell and electric vehicle markets are almost doubled owing to notable expansion in China, Japan, and Korea. The World Development strategy comparison of fuel cell vehicle and hydrogenation station (Table 1) depicts the current scenario and future development goals of the USA, Japan, China, and Europe.

1.1.1 Hydrogen Energy Scenario in the USA

Demand for hydrogen in the USA could reach around 22-41 million metric tons per year (MMT/yr) by 2050, approximately 10 MMT/yr at present. This would reduce US petroleum use by 15%, and 12 MMT/yr of hydrogen would fuel 18% of cars and 26% of light-duty trucks, with a further 5 MMT/yr fueling 22% of the medium- and heavy-duty vehicle [2]. Nearly all the hydrogen produced in the US is used for refining petroleum, treating metals, producing fertilizer, and processing foods.

1.1.2 Hydrogen Energy Scenario in Japan

In 2020, Japan declared that by 2050 Japan would reduce greenhouse gas emissions to net-zero and achieve a carbon-neutral, decarbonized society. Hydrogen is supposed to play a significant role in Japan's clean energy transformation. Japan was one of the first countries to launch a national hydrogen strategy to make hydrogen cost-effectiveness comparable to natural gas. As per their goals, by 2030, Japan aims to have more than 5 million residential fuel cells, 800,000 hydrogen fuel cell vehicles, and to set up a global hydrogen supply chain. This will result in providing valuable lessons to the international energy community. Japan is well-positioned to push for an internationally shared vision on making hydrogen an excellent energy source [3].

1.1.3 Hydrogen Energy Scenario in China

China is one of the world's largest hydrogen producers. They produce around 22 million tons of hydrogen per year, equivalent to one-third of the world's total. However, most of focus on using green hydrogen to power its fuel cells. Nevertheless, at present, only 3 % of China's hydrogen is made from renewable resources. China is considered a pioneer, and they are aiming for 5,000 fuel-cell vehicles by 2020 and 1 million fuel-cell vehicles by 2030 [4].

TABLE 1

World Development strategy comparison of fuel cell vehicle and hydrogenation station. [4]

Country/Region		Current Scenario	Future Development Goals			
		2018-2022	2025	2030	2040	2050
Japan	Hydrogen fuel cell vehicle	2,839	200,000	800,000	6,000,000	
	Hydrogenation station	113	320	900		
United States	Hydrogen fuel cell vehicle	5,899		1,000,000		
	Hydrogenation station	41		200		
EU	Hydrogen fuel cell vehicle	1,033	300,000	1,200,000		5,3000,000
	Hydrogenation station	152	1,500	3,700		15,000
South Korea	Hydrogen fuel cell vehicle	900-67,000			2,900,000	
	Hydrogenation station	14-310			1,200	
China	Hydrogen fuel cell vehicle	2,836	100,000	1,000,000		
	Hydrogenation station	61	350	1,000		

1.1.4 Hydrogen Energy Scenario in Europe

The EU consumes around 8 million tonnes of hydrogen every year, mainly produced from fossil gas rather than renewable power. Therefore Hydrogen technologies will play an essential role in reaching the Green Deal's 2050 net-zero targets by decarbonizing hard-to-abate sectors, according to a study by Hydrogen4EU.

The study states that allowing low-carbon hydrogen to contribute at its full potential and renewable hydrogen could save Europe over €2 trillion ($2.42 trillion) through 2050 [5].

1.1.5 Hydrogen Energy Scenario in South Korea

South Korea unveiled a hydrogen roadmap in 2019 that outlined a goal of producing 6.2 million FCEVs and rolling out at least 1,200 refilling stations by 2040 and announced "Green New Deal" in 2020, which includes investment and construction of three hydrogen cities by 2022, with a further three such cities to be added by 2025. These selected hydrogen cities are expected to provide hydrogen production and shipping infrastructure to serve residents as a primary energy source. Therefore, Korea plans to become Northeast Asia's largest hydrogen hub and further achieve carbon neutrality by 2050 [6]. "In the hydrogen transition, South Korea has strong positions in FCEVs and stationary fuel cells," law firm Clifford Chance said in their report.

1.2 Indian Hydrogen Energy Scenario

The Indian hydrogen market was valued at approximately ₹350 crores in 2017 and is projected to reach about ₹590 crores by 2025, growing at a Compound Annual Growth Rate (CAGR) of 6.3% from 2018 to 2025 [7]. Hydrogen is already used extensively in India, mainly as an industrial feedstock in creating ammonia-based fertilizers.

1.2.1 Methods Used for the Production of Hydrogen

Most hydrogen in India is produced through reforming methane, resulting in significant carbon dioxide emissions.

An alternative means of production is electrolysis, where water is split into its component using electricity. India claims to be one of the first large-scale alkaline electrolysis facilities globally, which produced hydrogen from electricity at the Nangal Facility from 1962.

1.2.2 Ministry of New and Renewable Energy

The Ministry of New and Renewable Energy (MNRE) had seen hydrogen as an area of strategic interest since 2006 when the first Hydrogen and Fuel Cell Roadmap was launched(MNRE, 2006).

More recently, in 2016, MNRE published a report laying out a comprehensive plan for increasing the R&D activity. This included significant funding for different electrolysis technologies and their integration with renewable electricity sources, which has strong potential in India, given the cost and availability of renewable electricity.

1.2.3 Investment Requirement in India

In terms of the investment requirements, if India is to deploy green hydrogen as a clean energy solution for crucial sectors, including transport, industry, and power, by 2050, this would require significant investment in electrolysis. The

potential scale of hydrogen use in India is vast, which increase 3 to 10 times by 2050 [8].

2. LITERATURE REVIEW

Production of hydrogen is possible in many ways from the spectrum of initial raw materials available. Presently, hydrogen is mainly produced by the steam reforming of natural gas, leading to massive emissions of greenhouse gases. Around 50% of the global demand for hydrogen is currently met via Steam Reforming Process(SPR), about 30% from reforming chemical industrial off-gases, 18% from coal gasification, 3.9% from water electrolysis, and 0.1% from other sources [11]. Plasma and Electrolyte processes yield a high efficiency for hydrogen production, but unfortunately, they are considered energy-intensive processes.

Steam reforming is one of the most widely used and comparatively cheaper processes for hydrogen production [12]. In the near future, biomass is anticipated to become the most viable renewable organic substitute for petroleum. Biomass is available from a variety of sources, such as animal wastes, municipal solid wastes, crop residues, short-rotation woody crops, agricultural wastes, sawdust, aquatic plants, short rotation herbaceous species (e.g., switchgrass), waste paper, corn, and many others. Another promising method of hydrogen production is Pyrolysis or co-pyrolysis. Raw organic material is gasified at a pressure of 0.1–0.5 MPa in the temperature range of 500–900°C [13]. An Important method for the production of hydrogen in the coming future would be the electrolysis of water; at present, approximately only 4% of hydrogen worldwide is produced by this process [10]. Photo electrolysis is one of the renewable ways of hydrogen production that used solar energy in the electrolysis of water, exhibiting promising efficiency and costs, although it is still in the phase of experimental development. Currently, it is the most effective and least expensive method of hydrogen production from renewable resources. Aqueous phase reforming (APR) is an under development technology to process oxygenated hydrocarbons or carbohydrates of renewable biomass resources to produce hydrogen [14, 15].

In the past, there have been some developments in this process that allowed more efficient and less expensive commercial units due to the changes in the material of reformer tubes, better control, and understanding of carbon limits. Better catalyst and process concepts developed over time.

3. MATERIALS AND METHODS

When we look at the various processes for hydrogen production, the question lies in the advancement of alternative technologies for hydrogen generation to those based on fossil fuels. However, it is not feasible to consider only the environmental aspect. The processes to be taken into account for the future must consider environmental concerns and the most favorable economics.

3.1 Biomass Gasification

Gasification is a process in which organic materials get converted at temperatures generally above 700°C, without combustion, with a controlled amount of oxygen and/or steams into CO, H, and CO.

Gasification technology is commonly used with biomass and coal as a fuel feedstock, and it is very mature and commercially used in many processes. It is a variation of Pyrolysis. Therefore, it is based upon partial oxidation of the feedstock material into a mixture of nitrogen, known as "producer gas," carbon monoxide, carbon dioxide, hydrogen, methane, higher hydrocarbons, carbon monoxide, and carbon dioxide [16].

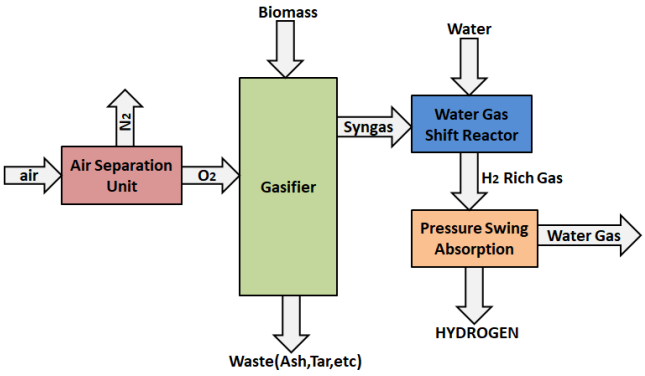


Fig-1: Biomass Gasification (Block Diagram)

3.1.1 Process

A controlled amount of oxygen or/and steam are added into CO, H, and CO. The carbon monoxide reacts with water to form carbon dioxide and more hydrogen via a water-gas shift reaction. The addition of steam or/and oxygen in the gasification process results in the production of "syngas" with an H2/CO ratio of 2/1, the latter used as feedstock to a Fischer-Tropsch reactor to make higher hydrocarbons (synthetic gasoline and diesel) or to a WGS reactor for hydrogen production.

Adsorber or special membranes can separate the hydrogen from this gas stream; the gas may also contain particulate matter, which is removed using cyclones and scrubbers. The particulate-free gas is compressed and then catalytically steam reformed to eliminate the tars and higher hydrocarbons. This is followed by high and low-temperature shift conversion reactions to produce additional hydrogen.

Finally, the hydrogen is separated from other products by pressure swing adsorption (PSA).

The main reactions taking place in biomass gasification are as follows:

$$\begin{matrix} Biomass + heat \longrightarrow H_2 + CO_2 + CH_3 + char \\ + \qquad\qquad\qquad\qquad + \\ steam \qquad\qquad CO + light\ and\ heavy\ HC \end{matrix} \quad (1)$$

However, the gasification process provides notable amounts of "tars" (a complex mixture of higher aromatic hydrocarbons) in the product gas even operated in the range of 800–1000°C [13]. A secondary reactor, which uses calcined dolomite or nickel catalyst, is used to clean and improve the quality of the product gas using the catalytic action [17].

For hydrogen production, a WGS process can be used to increase the hydrogen concentration, then followed by a separation process to produce pure hydrogen. Superheated steam (more than 900°C) has been used to reform dry biomass to achieve high hydrogen yields.

3.1.2 Challenges with this Approach

One major problem with this technology is that it must use a tremendous amount of resources to gather large amounts of biomass to the central processing plant.

This process typically suffers from low thermal efficiency since moisture in the biomass must also be vaporized. But, it can be performed in a fixed-bed or fluidized-bed reactor, with the latter reactor having typically better performance [17].

Ideally, gasification plants should use oxygen; however, the oxygen separation unit is cost-prohibitive for small-scale plants, limiting the gasifiers to the use of air, resulting in significant dilution of the product and the production of, Low-cost and efficient oxygen separators are required for this technology.

3.1.3 Advantages

Biomass is abundantly available as a renewable source of energy. It is less expensive than fossil fuels. It is clean carbon by-product fuel.

3.1.4 Applications

Currently, removing "tars" to acceptable levels for pure hydrogen production limits the commercialization of biomass-based hydrogen production. Cost-effective hydrogen production may require future development of smaller efficient distributed gasification plants for this technology for cost-effective hydrogen production. Mostly, gasification reactors are built on a large scale, have high logistics costs, and require massive amounts of material to be continuously fed. The efficiencies of the gasification plants are roughly around 70% [18].

3.2 Pyrolysis and Co-Pyrolysis

Pyrolysis is an endothermic thermal decomposition of biomass carried out in an inert atmosphere typically at 450 °C-550 °C. Gasification aims to produce gases, whereas Pyrolysis aims to produce bio-oil and charcoal.

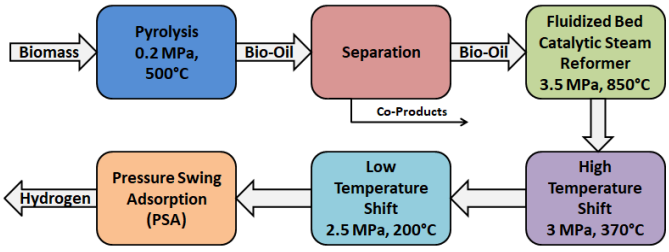


Fig-2: Pyrolysis Process (Block Diagram)

3.2.1 Process

Gasification aims to produce gases, whereas Pyrolysis aims to produce bio-oil and charcoal. Hydrogen can be made by reforming the biomass to a liquid bio-oil in a process called Pyrolysis.

The bio-oil so produced is a liquid composed of 85% oxygenated organics and 15% water. The bio-oil is then steam reformed in the presence of a nickel-based catalyst at 750 °C - 850 °C followed by shift conversion of CO and CO2. The following equation can generally describe the reaction:

$$Biomass \longrightarrow Bio.oil + char + gas \quad (2)$$
$$Bio.oil + H_2O \longrightarrow CO + H_2 \quad (3)$$
$$CO + H_20 \longrightarrow C0_2 + H_2 \quad (4)$$

Here equations (2), (3), (4) are stages of Pyrolysis, reforming, and shift, respectively. Based on the temperature ranges, pyrolysis processes are divided into low temperatures (up to 500°C), medium temperatures (500–800°C), and high temperatures (over 800°C). Fast Pyrolysis is one of the latest processes for transforming organic material into products having higher energy. The products of fast Pyrolysis appear in the entire phases formed (solid, liquid, and gaseous) [13].

3.2.2 Challenges with this Approach

The challenge while using this method of hydrogen production is the potential for fouling by the carbon formed, but some arguments claim that this can be minimized by appropriate design.

3.2.3 Advantages

Some of the Pyrolysis process advantages are relative simplicity, clean carbon by-product fuel flexibility, compactness, and reduction in carbon emissions.

Biomass is abundantly available as a renewable source of energy. It is less expensive than fossil fuels. It is clean carbon by-product fuel.

3.2.4 Applications

The application of co-pyrolysis has received interest in industrially advanced countries, as it should limit and lighten the burden of disposal of waste. Pyrolysis and co-pyrolysis are well-developed processes and could be used on a commercial scale.

Since CO and CO2 emissions are found out to be significantly low, and they can be operated in such that it recovers a significant amount of solid carbon, which is quickly sequestered, Pyrolysis may play an essential role in the future.

3.3 Aqueous Phase Reforming (APR)

Aqueous phase reforming produces hydrogen from biomass-derived oxygenated compounds such as sugar alcohols, glycerol, and sugars. APR is unique in that reforming is done in the liquid phase. The process generates hydrogen without volatilizing water, which represents significant energy savings.

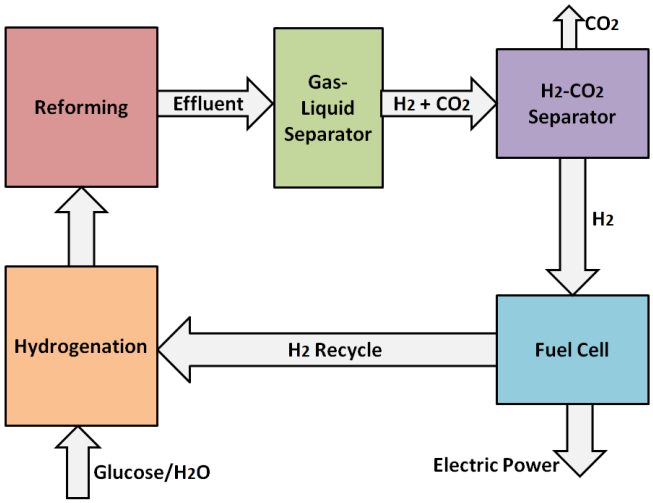


Fig-3: Aqueous Phase Reforming Process (Block Diagram)

3.3.1 Process

APR is a process where hydrocarbons or oxygenates are dissolved in water and react with water molecules in an aqueous phase at substantially lower temperatures (220–270°C) conventional alkane steam reforming (ca. 600°C) and high pressures to form H_2 and hydrocarbons. The low temperatures at which aqueous-phase reforming reactions occur minimize the undesirable decomposition reactions typically encountered when carbohydrates are heated to elevated temperatures. Aqueous feed concentrations of 10–60 wt% were reported for glucose and glycols [13].

APR is unique in itself as in this process, reforming is done in the liquid phase. The basic chemical reactions for APR are presented as:

$$C_nH_mO_n \longrightarrow nCO + mH_2 \quad (5)$$
$$CO + H_2O \longrightarrow CO_2 + H_2 \quad (6)$$

Equations (5) and equation (6) depict 'oxidation of oxygenated hydrocarbons' and 'WGS reaction'.

3.3.2 Challenges with this Approach

Catalyst selection is vital to avoid methanation, thermodynamically favorable, along with Fischer-Tropsch products, such as propane, butane, and hexane [14,15]. Most of the researches so far has been focused on supported Group 8 catalysts, with Pt containing solids having the highest catalytic activity. Even though they have comparatively lower activity, nickel-based catalysts have been evaluated due to nickel's low cost [14].

3.3.3 Advantages

It is possible to generate H_2 and CO_2 in a single reactor, as the water-gas shift reaction (WGS) is favorable at the same temperatures as in APR reactions, thus forming low amounts of CO. In contrast, the steam reforming processes require multistage or multiple reactors to achieve low levels of CO in the product gas.

It reduces the need to vaporize water, which provides a significant energy saving compared to conventional vapor-phase steam reforming processes.

3.3.4 Applications

Some researchers claim that this technology is more agreeable for converting biomass feedstock to hydrogen efficiently and selectively.

Recently, Rozmiarek reported an aqueous phase reformer-based APR, which achieved an efficiency of more than 55% with a feed composed of 60 wt.% glucose in water. However, the catalyst was not stable during long-term testing (200 days on stream) [23].

The reactors tend to be somewhat large due to moderate space-time yields. Improving catalyst activity and durability is a domain where significant developments can be made.

3.4 Electrolysis of Water

Hydrogen is produced by the electrolysis of water. In this process, electricity (electro-) is used to break down (negative-lysis) water (H_2O) into oxygen and hydrogen without any harmful or polluting side products. Fig. 4 depicts the Basic representation of the Electrolysis Process.

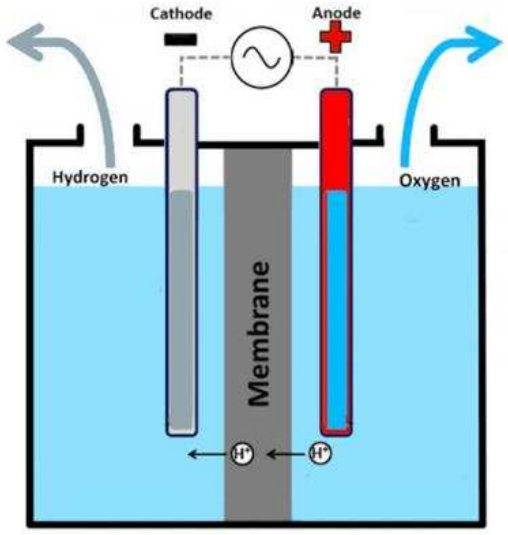


Fig-4: Basic Representation Electrolysis of Water Process

3.4.1 Process

Water electrolysis is when the chemical bonds present in the water molecule break into hydrogen and oxygen using a direct current passing through two electrodes in a water solution. The primary chemical reaction that takes place can be written as:

$$H_2O + Energy \longrightarrow {}^1/_2\, O_2 + H_2 \qquad (7)$$

At room temperature, the splitting of water is negligible, approximately 10-7 moles/liter, because pure water is a very poor conductor of electricity. Therefore, acid or base is added to improve the conductivity. Sulphuric acid is generally used as the electrolyte in water electrolysis, and the electrodes are of platinum (Pt), which does not react with sulphuric acid. For water electrolysis, the energy is required as electrical energy from a DC power source. In an alkaline electrolyzer, KOH, NaOH, and H2SO, solution is mainly used with water. The solution splits into ions positive and negative ions, and these ions readily conduct electricity in a water solution by flowing from one electrode to the other. Water electrolysis technology can be divided into three subcategories: Polymer/proton electrolyte membrane electrolysis, Alkaline electrolysis, and Steam electrolysis.

3.4.2 Challenges with this Approach

In comparison with the previous methods described, electrolysis is a highly energy-demanding technology.

3.4.3 Advantages

The chemical energy acquired per electrical energy supplied of the electrolysis of water in practice reaches 50–70% [19]. It is essentially converting electrical energy to chemical energy in hydrogen, with oxygen as a valuable by-product.

This process is ecologically very clean because no greenhouse gases are formed, and the oxygen produced has further industrial applications.

3.4.4 Applications

The most common application of electrolysis technology is alkaline-based electrolysis, but proton exchange membrane or polymer electrolyte membrane (PEM) and solid oxide electrolysis cells (SOEC) have also been developed [20, 21]. SOEC electrolysers are the most efficient but still are under-developed. SOEC technology faces difficulties due to corrosion, seals, thermal cycling, and chrome migration. PEM electrolysers are more efficient than alkaline and do not have corrosion problems and seals issues like SOEC; however, they cost more than alkaline systems. Alkaline systems are the most developed and have low capital costs. They have the lowest efficiency; therefore, they have the highest electrical energy cost [13].

3.5 Photo Electrolysis

Photoelectrolysis is also one of the renewable ways of hydrogen production in the phase of experimental development, exhibiting promising efficiency and costs, although it is still in the phase of experimental development. Currently, it is the most effective method and least expensive method of hydrogen production from renewable resources.

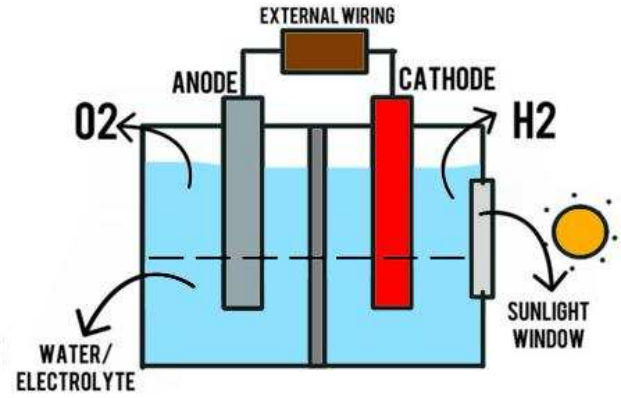


Fig-5: Basic Representation of Photo Electrolysis Process

3.5.1 Process

A photoelectrode is a semiconducting device that absorbs solar energy and creates the necessary voltage and current to decompose water molecules into hydrogen and oxygen. Photoelectrolysis employs a photoelectrochemical (PEC) light collection system for the electrolysis of water to take place. The semiconductor photoelectrode will generate enough electrical energy to support the induced reactions of hydrogen and oxygen if it is submerged in an aqueous electrolyte (exposed to solar radiation).

In the production of hydrogen, electrons are released into the electrolyte, while free electrons are required for the generation of oxygen. Photo-electrochemical (PEC) water splitting is one of the potential techniques for clean solar hydrogen production and has been utilized in small- to large-scale hydrogen generators.

There are three options for the arrangement of photoelectrodes in the assembly of PECS, Photo-anode made of n-type semiconductor and cathode made of metal, photo-anode made of n-type semiconductor, and photocathode made of p-type semiconductor, and Photo-cathode made of p-type semiconductor and anode made of metal.

3.5.2 Challenges with this Approach

The photovoltaic layer is produced from some of the light-absorbing semiconductor materials.

The light absorption of the semiconductor material is directly proportional to the performance of the photoelectrode. Semiconductors with wide bands provide the necessary potential for the splitting of water.

The photoelectrochemical cell is also influenced by the catalytic layers for the performance of the electrolysis and requires suitable catalysts for water splitting. The encased layer is another crucial component of the photoelectrode that prevents the semiconductor from rusting inside the aqueous electrolyte. This layer must be highly transparent to provide the maximum solar energy so that it could reach the photovoltaic semiconducting layer.

3.5.3 Advantages

The photoelectric device eliminates the need for a separate electrolyzer and power generator, which reduces the overall costs and increases the overall efficiency of the system. Photoelectrolysis systems are still in the experimental stage. However, Photoelectrolysis will become an essential means to future hydrogen production because it is powered by renewable solar energy [22]

This process is ecologically very clean because no greenhouse gases are formed, and the oxygen produced has further industrial applications.

3.5.4 Applications

Although This technology is in the experimental and modification stages, it already demonstrates promising efficiency and hydrogen production costs. The Non-renewable hydrogen production methods will only serve as a short-term supply for the hydrogen economy because they offset the benefits of fuel cell vehicles by consuming fossil fuels and emitting greenhouse gases; therefore, methods like Photoelectrolysis are worth investing in.

3.6 Thermo Water Splitting

Thermochemical cycles have been developed since the 1970s and 1980s when they searched for new sources to produce alternative fuels during the petroleum crisis. Hydrogen is widely considered the fuel in the power of nonpolluting vehicles, domestic heating, and aircraft.

The abundant supply of water and sunlight offers us an affordable alternative source to produce hydrogen apart from fossil fuels and biomass.

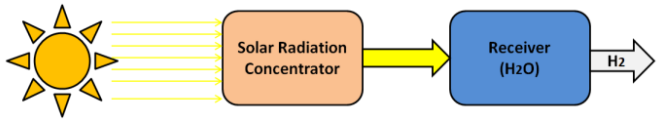


Fig-6: Thermo Water Splitting Process (Block Diagram)

3.6.1 Process

Thermochemical water splitting processes use high-temperature heat (500-2000 °C) to drive a series of chemical reactions that produce hydrogen. The chemicals used in the process are reused within each cycle, creating a closed-loop that consumes only water and produces hydrogen and oxygen. The thermal dissociation of water is described as follows:

$$H_2O + heat \longrightarrow {}^1/_2\, O_2 + H_2 \qquad (8)$$

3.6.2 Challenges with this Approach

An effective technique is needed to separate oxygen and hydrogen to avoid ending up with an explosive mixture. Semi-permeable membranes based on zirconium oxide and other high-temperature materials can be used for this purpose. If the product gas mixture is quenched to lower temperatures, separation can also be achieved. Then, Palladium membranes can be used for effective hydrogen separation.

A material, which satisfies high bandgap and corrosion-resistant, is not available commercially. Therefore, there is a need to process such material. Photo-electrodes should be resistant to undesired reactions.

3.6.3 Advantages

PEC technology depends on solar energy, which is a renewable source.

Photo-voltages are generated on both electrodes in the bi-photo-electrode PEC system interface, resulting in the formation of an overall photo-voltage that is sufficient for water decomposition without the application of a bias. It is environmentally safe, with no harmful by-products. It may be used on both large and small scales.

4. ECONOMICAL ASPECT

In the present scenario, the cheapest and most widely used method for hydrogen production is the steam reforming of methane (natural gas). This method produces about half of the world's hydrogen production, with the cost of production about 7 USD/GJ. Partial oxidation of hydrocarbons provides a comparable price for hydrogen. However, the emission of greenhouse gases generated by these processes must be purified and stored, increasing the hydrogen price by 25–30% must be evaluated [13].

Thus the renewable energy seeks attention that includes thermochemical processes like gasification and Pyrolysis of biomass. The price of hydrogen production from gasification of biomass ranges from 10–14 USD/GJ and that from pyrolysis 8.9–15.5 USD/GJ. The cost of hydrogen obtained is about two to three times greater than the price of hydrogen obtained by the SR process. The final cost of production primarily depends on the equipment, availability, and cost of feedstock [9].

Electrolytic processes can be termed as highly effective when it comes to greener methods of production. It is one of the simplest technologies for producing hydrogen without by-products. Meanwhile, the input electricity cost plays a key role in the price of hydrogen obtained and is relatively high. According to the researchers, "the dominant methods for hydrogen production by the year 2030 will be 3716 catalyzed biomass gasification and steam reforming of natural gas. The use of solar energy in the given scenario is also a possibility. The role of solar energy is predicted to increase by 2050 [9].

5. CONCLUSIONS

Hydrogen is widely considered the fuel for the power of nonpolluting vehicles, domestic heating, and for aircraft. The abundant supply of water and sunlight offers us an affordable alternative source to produce hydrogen apart from fossil fuels and biomass**.** The development of hydrogen (H2) generation technologies is the topic on which a tremendous amount of research is going on. At present, the most developed and used technology is the reforming of hydrocarbons. Also, Significant developments in other hydrogen generation technologies from renewable resources such as biomass, water, and solar are the current focus to decrease the dependence on fossil fuels. Many processes are under development with minimal environmental impact.

Hydrogen is produced from various feedstock available almost everywhere. These technologies may reduce the world's dependence on fuels that comes primarily from unstable regions. The increase of hydrogen production within the country might increase economic security and national energy. The ability of hydrogen to be produced from a wide variety of feedstock and using various processes may make every region of the world able to have hydrogen energy. Hydrogen may prove to be the most ubiquitous fuel available.

REFERENCES

[1] Hydrogen Council, "Hydrogen, Scaling Up". November 2017

[2] Ruth, Mark, Paige Jadun, Nicholas Gilroy, Elizabeth Connelly, Richard Boardman, A.J. Simon, Amgad Elgowainy, and Jarett Zuboy. 2020. "The Technical and Economic Potential of the H2-Scale Concept within the USA". Golden, CO: National Renewable Energy Laboratory. NREL/TP- 6A20-77610. https://www.nrel.gov/docs/fy21osti/77610.pdf.

[3] International Energy Agency. Japan 2021 Energy Policy Review. 2021. IEA. https://www.iea.org/reports/japan-2021.

[4] Meng X et al., Status quo of China hydrogen strategy in the field of transportation and international comparisons, International Journal of Hydrogen Energy, https://doi.org/10.1016/j.ijhydene.2020.11.049

[5] Kate Abnett. "Six European Union Countries Lead Push for Clean Hydrogen Support.", Reuters, 2020. https://www.reuters.com/article/us-eu-energy-hydrogen-idUSKBN23M18G.

[6] Bernadette Lee. South Korea Plans Northeast Asia's Largest Hydrogen Hub. 2021. IHS. Markit. https://ihsmarkit.com/research-analysis/south-korea-plans-northeast-asias-largest-hydrogen-hub.html.

[7] India Hydrogen Market by Delivery Mode (Captive and Merchant), Technology (Steam Methane Reforming, Partial Oxidation of Oil, Coal Gasification, Methanol Reforming, Ammonia Cracking, and Electrolysis of Water), and End User (Chemical, Petroleum Refining, Metal Processing, Glass Industry, Edible Fats and Oils, Energy, and Others) - Global Opportunity Analysis and Industry Forecast, 2018-2025 at: https://www.alliedmarketresearch.com/india-hydrogen-market accessed on 09/05/2021

[8] The Energy and Resources Institute. Make Hydrogen in India. 2020. TERI. https://www.teriin.org/sites/default/files/2020-06/Hydrogen-Policy-Brief.pdf.

[9] Mustafa Balat and Mehmat Balat, Political, Economic and Environmental Impacts of Biomass Based Hydrogen, International Journal of Hydrogen Energy, https://doi.org/10.1016/j.ijhydene.2009.02.067

[10] A. Konieczn et al., Catalyst development for thermocatalytic decomposition of methane to hydrogen, International Journal of Hydrogen Energy, https://doi.org/10.1016/j.ijhydene.2007.07.054

[11] N. Muradov and T. Veziroğlu, "From Hydrocarbon to Hydrogen-Carbon to Hydrogen Economy," International Journal of Hydrogen Energy, https://doi.org/10.1016/j.ijhydene.2004.03.033

[12] Joan M Ogden, Margaret M Steinbugler, Thomas G Kreutz, A comparison of hydrogen, methanol and gasoline as fuels for fuel cell vehicles: implications for vehicle design and infrastructure development, Journal of Power Sources, Volume 79, Issue 2, Pages 143-168, 1999, https://doi.org/10.1016/S0378-7753(99)00057-9

[13] Christos M. Kalamaras and Angelos M. Efstathiou, Hydrogen Production Technologies: Current State and Future Developments, 2013, https://www.hindawi.com/journals/cpis/2013/690627

[14] George W. Huber, et al., " Aqueous-phase reforming of ethylene glycol over supported Pt and Pd bimetallic catalysts", Applied Catalysis B Environmental, Volume 62, Issues 3–4, 2006, Pages 226-235, https://doi.org/10.1016/j.apcatb.2005.07.010

[15] Huber, G.W. and Dumesic, J.A. "An Overview of Aqueous-Phase Catalytic Processes for Production of Hydrogen and Alkanes in a Biorefinery", Catalysis Today, 111, 119-132, 2006. http://dx.doi.org/10.1016/j.cattod.2005.10.010

[16] Office of Energy Efficiency & Renewable Energy, "Hydrogen Production: Biomass Gasification", US Dept. Of Energy, https://www.energy.gov/eere/fuelcells/hydrogen-production-biomass-gasification

[17] Mohammed Asadullah et al., "Novel biomass gasification method with high efficiency: catalytic gasification at low temperature", Green Chemistry, Issue 4, pp. 385-389, 2002, https://doi.org/10.1039/B204536M

[18] Woranee Paengjuntuek, Jirasak Boonmak, Jitti Mungkalasiri, Energy Efficiency Analysis in an Integrated Biomass Gasification Fuel Cell System, Energy Procedia, Volume 79, Pages 430-435, 2015, https://doi.org/10.1016/j.egypro.2015.11.514

[19] Turner, J. Et al., "Renewable hydrogen production", Int. J. Energy Res., 32: 379-407, 2008, https://doi.org/10.1002/er.1372.

[20] S.A. Grigoriev, V.I. Porembsky, V.N. Fateev, Pure hydrogen production by PEM electrolysis for hydrogen energy, International Journal of Hydrogen Energy, Volume 31, Issue 2, Pages 171-175, 2006, https://doi.org/10.1016/j.ijhydene.2005.04.038

[21] J. Pettersson, B. Ramsey, D. Harrison, A review of the latest developments in electrodes for unitised regenerative polymer electrolyte fuel cells, Journal of Power Sources, Volume 157, Issue 1, Pages 28-34, 2006, https://doi.org/10.1016/j.jpowsour.2006.01.059

[22] Robret Huang, Hydrogen production by Photoelectrolysis Accessed on 10/05/2021 at: https://bhuang02.tripod.com/photoelectrolysis.htm .

[23] B. Rozmiarek, "Hydrogen generation from biomass-derived carbohydrates via aqueous phase reforming process", Department of Energy USA, 2008.

Feature Enhanced Capsule Networks for Robust Automatic Essay Scoring

Arushi Sharma[1,3⋆][0000−0003−4770−4577], Anubha Kabra[2,3⋆][0000−0003−1874−7711], and Rajiv Kapoor[3][0000−0003−3020−1455]✉

[1] Optum Global Advantage - arushi.net@outlook.com
[2] Adobe Systems - anubhakabradtu@gmail.com
[3] Delhi Technological University rajivkapoor@dce.ac.in

Abstract. Automatic Essay Scoring (AES) Engines have gained popularity amongst a multitude of institutions for scoring test-taker's responses and therefore witnessed rising demand in recent times. However, several studies have demonstrated that the adversarial attacks severely hamper existing state-of-the-art AES Engines' performance. As a result, we propose a robust architecture for AES systems that leverages Capsule Neural Networks, contextual BERT-based text representation, and key textually extracted features. This end-to-end pipeline captures semantics, coherence, and organizational structure along with fundamental rule-based features such as grammatical and spelling errors. The proposed method is validated by extensive experimentation and comparison with the state-of-the-art baseline models. Our results demonstrate that this approach performs significantly better on 6 out of 8 prompts on the Automated Student Assessment Prize (ASAP) dataset. In addition, it shows an overall best performance with a Quadratic Weighted Kappa (QWK) metric of 81%. Moreover, we empirically demonstrate that it is successful in identifying adversarial responses and scoring them lower.

Keywords: Automatic Scoring · Capsule Neural Networks · Adversarial Testing · BERT · Machine Learning

1 Introduction

Writing compositions have been widely adopted by all language proficiency exams. The manual evaluation process is taxing and laborious; hence globally standardized exams such as GRE [2] resort to automatic scoring systems. However, many state-of-the-art AES tools suffer from adversarial attacks [10, 19]. As a result, there is a dire need for computerized essay scoring systems that provide quick results while maintaining objectivity and accuracy in evaluations [17]. AES is a complex problem as these systems aim not just to point out grammatical or spelling errors but also to consider the semantics, identify the coherence in discourse, and ensure that the response is relevant to the question.

State-of-the-art AES systems can be broadly categorized into two types, feature engineered models and end-to-end deep learning models [14]. The feature-engineered

⋆ Equal Contribution - work done in Delhi Technological University

models use handcrafted surface-level features, such as the length of the essay and grammatical errors, for scoring the responses [11]. While they are easily explainable and modifiable with scoring criteria, they lack the understanding of response as a whole. They cannot mimic pattern organization or coherence based on word and sentence level relations [21]. The latter type of AES system explores extracting semantic relationships within an essay. This method relies on word embeddings, which are used to express response essays in low dimensional vectors followed by dense, CNN or LSTM layers to represent the semantics in the text [23, 1]. Models that incorporate deep learning techniques and handcrafted features aim to capture all the aspects required for essay scoring. However, they are vulnerable to adversarial attacks [15]. One of the challenges in such a task is identifying if the responses correctly answer the question. Consider an example of a question asking the test takers to describe an incident where they had to be patient. (*Prompt 7 Question: Write a story about a time when you were patient.*) If the test taker answers general sentences about patience and does not provide a personal story, the response should be scored lower as it does not answer the question. As shown in Table 1, the incorrect answer has all the features required for a good answer: correct grammar and spellings, topic relevance and coherence; yet this response does not qualify as the correct answer. While a human annotator can quickly identify this, an AES system requires learning each essay's underlying inconsistencies. Hence, understanding each essay response's composition, structure, and hierarchies need to be further explored to ensure correct scoring.

Table 1. For a given question: *Write a story about a time when you were patient*, the Table lists a correct and incorrect answer to illustrate that even a well-formed answer is marked incorrect simply because it does not answer the question.

Incorrect Answer	Correct Answer
Patience is the ability to endure difficult circumstances. Patience may involve perseverance in the face of delay. Patience is the level of endurance one can have before disrespect. It is also used to refer to the character trait of being steadfast...	I am not a patience person, as I can't sit in a room for more than five minutes, but there was one time I was patient, during my wife's operation. I was sitting quietly in the hospital. At that moment I felt that I needed that time to pray for her well being...

The problem above is addressed in this work, using the newly introduced Capsule Neural Network [24]. The CapsNet was introduced in image classification to overcome the drawbacks of a CNN network by allowing an effective combination of low-level features of images to high-level depiction using iterative routing. CNNs not only face problems in learning the transformational invariance of images but also are unable to ensure a local agreement between the features (due to the max-pooling layers). CapsNet understands the spatial correlation between a part and a whole and analyzes the current reference frame to generalize to new or unseen frames.

We aim to use this method to understand each response based on the intrinsic spatial relationship between the parts (sentences) and the whole (essay response). Moreover, this technique provides an understanding of the general structure of the response [35].

Hence, the organization and transition amongst the lines in the response will be under scrutiny. We did not shy away from using key handcrafted features to enhance the model's capability to score the responses. We call the CapsNet architecture, CapsRater, and the feature engineering enhancement, FeatureCapture (Figure 1). To provide a contextualized and semantic initial representation of essay responses, we use the BERT embeddings [9]. This architecture improves performance on six out of eight prompts and increases the overall average QWK score. Moreover, to check the robustness of this work, we extensively experimented with adversarially crafted essay responses [19]. CapsRater + FeatureCapture successfully detected these adversarial attacks and scored them relatively lower than the original essay responses.

Following are the significant contributions of this work[4]:

- We provide a novel architecture with two key components: BERT enhanced CapsNet model and the feature extraction model for Automatic Essay Scoring. This is the first work that leverages Capsule Neural Networks for Automatic Essay Scoring to the best of our knowledge.
- We develop critical features with emphasis on official rubrics for scoring. We empirically demonstrate the importance of these features in improving the QWK.
- The proposed method shows higher scores on the QWK metric than state-of-the-art models with a boosted overall average QWK of 81%.
- We analyze the proposed work's performance on multiple adversarial attacks proving that our model is considerably robust. It successfully scores most of the adversarial text lower.

The paper's organisation is as follows: we discuss related work in section 2, and the proposed pipeline in section 3. We illustrate substantial experimentation in section 4, to evaluate performance with baseline models. In the same section, we also assess our model's sturdiness on the adversarially perturbed datasets.

2 Related Work

The previous work on essay scoring relies on human experts who build domain-specific features to check the lexical and grammatical errors. The systems then employ machine learning classifiers to predict the essay scores. For example, works such as, [20, 3, 32, 6, 16], trained Naive Bayes, Linear Regression and Rank Support Vector Machine (RankSVM) models for essay scoring task. EASE [11], a popular AES engine, applies text analysis and feature engineering to several regression models. Using handcrafted features has had immense success on the AES task. While there are simple features such as sentence length, word count, there are also other features that have a convex mechanism of engineerings, such as readability [33], textual and discourse coherence [5, 26]. Incorporating each of these requires specialized focus and a domain-specific approach.

More recent approaches have turned to neural networks, which encode an essay into richer representations of low-dimensional embeddings. For example, Alikaniotis et al.

[4] Our code is available at : https://github.com/ECMLPKDD/CapsRater-FeatureCapture

[1] applied a deep LSTM layer, Taghipour et al. [27] employed the CNN layer. They map texts into sequences that can account for the variable input lengths. These strategies are popular because they provide an end-to-end solution. A fundamental limitation of deep learning approaches in this domain is that they are susceptible to adversarial attacks [10]. Researchers have conducted a study to showcase how simple tricks can deceive state-of-the-art AES tools [19]. In this light, Farag et al. [12] applied window-based local coherence to catch adversarial attacks. However, detecting and mitigating advanced attacks is still an open problem.

Previous works on Capsule Neural Networks showcased promise on the image, as well as text classification tasks citehinton2011transforming [35]. The capsules utilize all the feature information, therefore, address the issues with information loss in CNNs. Variants of Capsule architectures have recently experimented with in text classification applications. For example, Zhao et al. [35] proposed two models to stabilize the routing mechanism between capsules from disturbances in the text, such as stopwords, as these do not contribute to the classification task at hand. Kim et al. [18] employ a static routing mechanism in place of the dynamic one, based on the observation that the document semantics can remain the same with a different order of sentences. Saha et al. [25] studied the joint optimization capabilities of BERT and capsule layer in their classification task.

Due to the success of Capsule Neural Networks to capture spatial inter-dependencies in text [36], we employed the CapsNet for the AES task. Our work leverages BERT embeddings to utilize the pre-trained contextual features. The capsule architecture captures the semantic inter-word relationships and spatial patterns of words and transfer learned parameters. It can effectively encode the information required for essay evaluation and learn the complex patterns in the data.

3 Methodology

Our architecture consists of two independent pipelines, called CapsRater and Feature-Capture. We report results on both of these models. However, the best performing model is their combination (Figure 1), as shown in section 5. To combine these models, we took the mean over their class-wise probabilities and passed the output through a final dense layer to get the resultant score vector.

3.1 CapsRater

Capsule Neural Networks were introduced to accurately identify hierarchical relationships between objects and the features that constitute those objects. A capsule is a structure that essentially contains information about the probability and orientation of these features. Following [24], we use the capsules to attain a vector output from the feature detectors in CNN. The capsule representation is used in place of the pooling layer's scalar output, which discards the text's positional information.

We employed BERT embeddings of essay text and fine-tuned them using the Capsule framework. BERT provides contextual word representations by applying bidirectional training of transformer to language modelling. During the pre-training on the

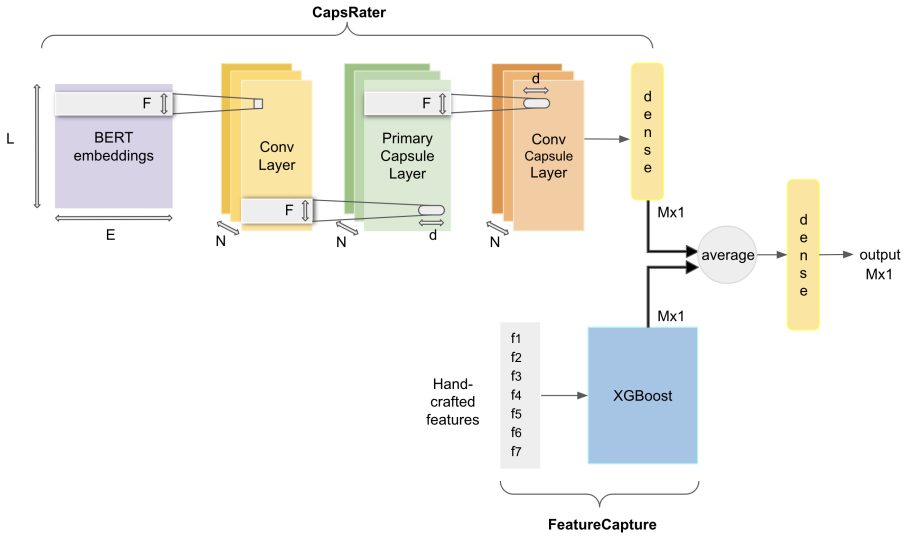


Fig. 1. The combined architecture of CapsRater + FeatureCapture model.

Masked Language Model task, BERT's architecture enables it to capture the entire text instead of sequential reading of directional models. As a result, BERT has a deeper understanding of language context and flow [9]. Thus, using BERT representation will provide the non-linear neural layers with the richer vector representations for the essay responses.

Consider the input sentence, S of length L, forms E dimensional BERT embeddings; then the input is represented as $S \in R^{L \times E}$.

Convolution Layer: We passed the input through the convolution layer to extract local textual features. In this step, the convolutional filter matrix, $W^a \in R^{F \times V}$ (F is the size of filter), slides across input of $s_{i:i+F-1}$ dimensions, multiplies with the input element-wise, and generates feature maps of size, $k^a \in R^{L-F+1}$. The resultant is added to a bias term, b_0, and a ReLU activation function, f, is applied to it. Mathematically, this is denoted as:

$$k_i^a = f\left(\mathbf{s}_{i:i+F-1} \circ W^a + \mathbf{b}_0\right) \tag{1}$$

The above operation is iterated for N filters to widen the feature extraction process and concatenate the output [35]. Therefore there are N feature maps generated, represented by $K = [k_1, k_2, ..., k_N] \in R^{(L-F+1) \times N}$

Primary Capsule Layer: The feature maps are then passed through a convolution operation, with filter, W^b, forming the first layer of capsules, denoted as $\Theta \in R^{(L-F+1) \times d}$, ($d$ is the dimension of capsule). This layer preserves the initial parameters belonging to each input feature instead of the scalar output from CNN's pooling operations. As a

result, the capsules contain more information about the input, given as:

$$\Theta_i = (W^b)^T \times K_i \ + b_1 \tag{2}$$

where i denotes the size of the filter matrix: $1 \rightarrow L - F + 1$ and b_1 is the bias term.

A nonlinear squash function is applied on the above vectors.

$$\Theta_i = \frac{\|\Theta_i\|^2}{1 + \|\Theta_i\|^2} \frac{\Theta_i}{\|\Theta_i\|} \tag{3}$$

We performed the above steps with N number of filters and concatenated their corresponding outputs. For our domain, the capsule layer is aware of the semantics and ordering of sentences due to the vector representation of the instantiated parameters [35].

Part-Whole Relationship: Hinton et al. [13] defines this step as assigning part-to-whole.

We made use of two levels of capsules; each lower layer capsule assigns a vote vector $V_{low \rightarrow up}$ to each upper layer capsule. These vectors represent how much information is transferred from the different input capsules to the respective output capsule.

$$V_{low \rightarrow up} = W_{low \rightarrow up} \times \Theta_{low} \tag{4}$$

where $W_{low \rightarrow up}$ is a transformation matrix. These matrices solve exponentially taxing convolutions, and their insubstantial representation [24]. Moreover, they provide automated learning of part-whole relationship.

Dynamic Routing: This is an iterative process that builds a non-linear map to ascertain that each lower capsule's output is matched to the suitable upper capsule in the next layer. It controls the connection between the higher and lower layers' capsules. Following [24], there is an assignment probability $C_{low \rightarrow up}$ associated with input capsules, which measures the similarity between vote vector and output capsule. It is calculated as the multiplication of probability of the upper capsule for each lower capsule, $A^{'}_{low \rightarrow up}$, with the softmax of logits of the assignment probability, $B_{low \rightarrow up}$.

$$C_{low \rightarrow up} = A^{'}_{low \rightarrow up} \times \frac{\exp\left(B_{low \rightarrow up}\right)}{\sum_{up=1}^{N} \exp\left(B_{low \rightarrow up}\right)} \tag{5}$$

Here $B_{low \rightarrow up}$ is initialized as all 0s, and it measures the proportion of input capsule that makes the output capsule. It is updated according to the agreement between the upper layer capsule and the vote vector.

$$B_{low \rightarrow up} = B_{low \rightarrow up} + \Theta_{up}.V_{low \rightarrow up} \tag{6}$$

Output capsules are formed from the weighted sum of vote-vectors.

$$\Theta_{up} = squash(\frac{\sum_{low=1}^{H} (C_{low \rightarrow up} V_{low \rightarrow up})}{\sum_{low=1}^{H} (C_{low \rightarrow up})}) \tag{7}$$

$$A_{up} = |\Theta_{up}| \tag{8}$$

here, A_{up} is the activation probability of the upper layer capsule.

Convolutional Capsule Layer: The next layer gives the upper layer capsules formed from multiplying transformation matrices with the lower capsules, followed by dynamic routing. Using this routing mechanism, the capsules capture the importance and coherence of words while leaving out nonessential information. Finally, the upper layer capsules are flattened and passed through a dense layer.

3.2 FeatureCapture

Table 2. The key features extracted from the text to pass through the FeatureCapture model.

S. No.	Type	Feature
1	Prompt Based	Number of words
2		Number of ! or ?
3		Correct POS tags
4		Number of spelling errors
5		Number of grammatical errors
6	Similarity Based	Between prompt and responses
7		Between response sentences

Handcrafted features play an essential role in enhancing the performance of AES systems [21]. This can be observed in CapsRater + FeatureCapture model scores in Table 4, where adding features has illustrated best scores for the majority (5 out of 8) of prompts in the ASAP dataset. In this work, we used two types of handcrafted features: prompt-based and similarity-based (listed in Table 2). These were empirically decided by manual inspection of the responses and official rubrics given in the dataset. The prompt-based features are inspired by the EASE system [11].

- **Prompt Based Features**: We observed that shorter essays are scored lower by the annotators. Therefore, *Number of words* is an important length-based feature. We also took the frequency of *sentences ending with ! or ?* into account, as the essay response should not be primarily constituted of exclamatory or questioning sentences. For the next feature, *Correct POS tags*, we counted the number of erroneous unigrams and bigrams and subtracted those from total unigrams and bigrams. Using this, we get a statistical understanding of how coherent the essay response is. The above features were defined using the NLTK[5] library. Essays with many grammatical and spelling errors are penalized higher by the annotators; hence we included these in the feature space. *Number of spelling errors* were counted using pyspellchecker [4]. It relies on the Levenshtein Distance algorithm and compares all permutations in the frequency word list to correct an incorrect spelling. For identifying the *Grammatical Errors*, we use the popular grammar checking application called LanguageTool's python wrapper [29].
- **Similarity Based Features**: *The similarity score between the prompt and essay* shows if the response essay has borrowed text from the prompt by comparing the similarity between them. *Similarity between the sentences* gives an analysis of the amount of repetition within a response. Higher repetition reported by this feature should be scored lower. For calculating both, we use a fuzzy matching library, called rapidfuzz [22].

[5] https://www.nltk.org/

All the above-extracted features are normalized to the 0-1 range, stacked and passed into the XGBoost [7], which is an ensemble learning method. Multiple individual base learners (models) are trained and combined for a final prediction. XGBoost has base learners that may generate average performance, but when sequentially added, they rectify the errors and lead to efficient predictions. Moreover, this algorithm is immensely scalable, and it relies on distributed computing which enhances fast learning [7]. The output of XGBoost is a probability vector for each target class. The final output vectors from CapsRater and FeatureCapture are averaged and passed through a dense layer for the final result.

4 Experimentation

4.1 Dataset

We perform our experiments on the widely used and accepted dataset for AES tasks, namely Automated Student Assessment Prize (ASAP). The dataset comprises 8 prompt questions, which students of grades 7 to 10 answer. The total number of answered essays are 12,976. More details about the dataset are shown below in Table 3.

Table 3. Description of the ASAP-AES Dataset used for evaluation of AES systems. Here RC refers to Reading Comprehension and # represents the count.

Prompt	# Responses	Type	Avg # words	Avg # sentences	Score Range
1	1783	Argumentative	350	23	2-12
2	1800	Argumentative	350	20	1-6
3	1726	RC	150	6	0-3
4	1772	RC	150	4	0-3
5	1805	RC	150	7	0-4
6	1800	RC	150	8	0-4
7	1569	Narrative	250	12	0-30
8	723	Narrative	650	35	0-60

4.2 Evaluation Metric

We use the Quadratic Weighted Kappa (QWK) metrics for evaluation. It is a commonly used and accepted metric [27, 21] which measures agreement between the AES scorer and the human annotators. QWK is calculated as:

$$k = 1 - \frac{\Sigma_{ij} w_{ij} Obs_{ij}}{\Sigma_{ij} w_{ij} Exp_{ij}} \tag{9}$$

Here, Obs and Exp are the observed and expected scores matrix respectively, while w denotes the weights. The scores assigned by the human and machine graders are i and j respectively. Given N is the number of possible scores, the weight matrix is defined as:

$$w_{ij} = \frac{(i-j)^2}{(N-1)^2} \tag{10}$$

The range of QWK score is from 0 to 1. The higher the score, the closer the machine-human agreement.

4.3 Baselines

We compared the proposed work with the recent state-of-the-art baselines: *EASE* [11], developed by *EdX*, is a feature-based model relying on n-grams and prompt word overlap. It performed third-best in the ASAP-AES competition. Taghipour et al. [27] build an architecture that applies CNN and LSTM with a mean over time layers. This model performs better than *EASE* on the ASAP dataset. *HISK+BOSWE* [8] is a statistical technique that captures text-based features using string kernels along with word embeddings. A reinforcement learning methodology, called *RL1*, was proposed by Wang et al. [30]. They used the QWK score as the reward function, which is governed by positive or negative feedback. *SkipFlow* [28] proposes the *neural coherence features* that capture semantics and coherence using deep neural networks. It uses the hidden states for extracting more information about the formation of the response. *MemoryNetworks* [34] takes one sample essay belonging to each score in the score range and saves it in the memory. This sample essay is used to calculate the similarity with new essays to score them. Using memory networks in grading helps them boost performance on 7 out of 8 prompts. *TSLF* [21] uses feature engineering along with neural networks for scoring. They form features for coherence, semantics and relevance using BERT embedding and employ the SVM classifier. They perform relatively higher than the state of the art methods. However, Kumar et al. [19] have shown that they lack robustness to tackle all kinds of adversarial attacks. *R^2BERT* [31] has multiple objective approaches where they explored two loss functions: the mean squared error and the batch-wise ListNet loss. They report improved results on baselines.

4.4 Implementation

We perform prompt-wise training on our models. The prompts in the dataset vary in terms of the genres (Argumentative, Reading Comprehension and Narrative), the scoring rubrics, and the grade of study of test-takers. We load the response essays to pass them through CapsRater and FeatureCapture, respectively. The first step in CapsRater involves transforming the data into the BERT-base model's embeddings ($V = 768$). It is then passed through the CapsRater model, where we use 32 filters, of size $F = 3$ words, in the first and second convolution layers while 16 filters in the convolution capsule layer. The capsules are also set as 16-dimensional vectors, and the dynamic routing process is iterated thrice for optimum loss convergence. In FeatureCapture, we use the XGBClassifier with the gbtree booster method. The max_depth parameter is set to 6, the objective function to multi:softprob, and n_estimators to 1000. The learning rate is set to 1e-4. For the CapsRater pipeline, we used the standard, cross-entropy loss function, Adam optimization, and the model is trained for 50 epochs with a 1e-4 learning rate.

5 Result and Analysis

This section compares CapsRater, FeatureCapture, and their combined pipeline with the prominent baseline works on the ASAP dataset. Table 4 shows the QWK scores on all 8 prompts available in the data. Moreover, we provide the average score to represent the overall performance. Table 4 reports that the proposed architecture boosts performance on six out of eight prompts, including both Argumentative, three Reading Comprehension and one Narrative prompt. Moreover, it leads to an overall increase of 2% in scores compared to the baselines.

Analysis: We had a thorough look at the essay responses and their scoring rubrics provided with the dataset. The rubrics can be categorized into three main parts[6]: Ideas + Content, Organization and Style. Ideas + Content focus on topic relevance and scrutinizes the main idea of

[6] https://www.kaggle.com/c/asap-aes

Table 4. Performance comparison of proposed work with baseline models for each prompt in the dataset. Scores are calculated using the QWK metric. The best performance for each prompt is **emboldened**

PROMPT	1	2	3	4	5	6	7	8	Avg
EASE (SVR) [11]	0.781	0.621	0.630	0.749	0.782	0.771	0.727	0.534	0.699
EASE (BLRR) [11]	0.761	0.606	0.621	0.742	0.784	0.775	0.730	0.617	0.705
CNN [27]	0.804	0.656	0.637	0.762	0.752	0.765	0.750	0.680	0.726
LSTM [27]	0.808	0.697	0.689	0.805	0.818	0.827	0.811	0.598	0.756
CNN+LSTM [27]	0.821	0.688	0.694	0.805	0.807	0.819	0.808	0.644	0.761
HISK+BOSWE [8]	0.845	0.729	0.684	0.829	0.833	0.830	0.804	0.729	0.785
RL1 [30]	0.766	0.659	0.688	0.778	0.805	0.791	0.760	0.545	0.724
SkipFlow [28]	0.832	0.684	0.695	0.788	0.815	0.810	0.800	0.697	0.764
MemoryNets [34]	0.830	0.720	0.720	0.820	0.830	0.830	0.790	0.680	0.780
TSLF [21]	0.852	0.736	0.731	0.801	0.823	0.792	0.762	0.684	0.773
R^2BERT [31]	0.817	0.719	0.698	0.845	**0.841**	0.847	0.839	**0.744**	0.794
CapsRater (CR)	0.852	0.750	0.743	**0.847**	0.800	0.780	0.812	0.702	0.785
FeatureCapture (FC)	0.791	0.677	0.693	0.818	0.782	0.771	0.762	0.699	0.749
CR + FC	**0.866**	**0.764**	**0.751**	0.844	0.837	**0.852**	**0.843**	0.715	**0.809**

the response. Organization refers to the logical and structural organization of the response. Each response should give a meaningful, cohesive and complete meaning. Finally, the style includes features that emphasize spelling and grammatical errors. It castigates repetitive word and sentence usage. We also observed that poor-scored essays were generally shorter. While CapsRater has strong capabilities of demonstrating high performance, combining with FeatureCapture exhibits faster convergence, as it adheres to the prompt rubrics. This is evident in Figure 2.

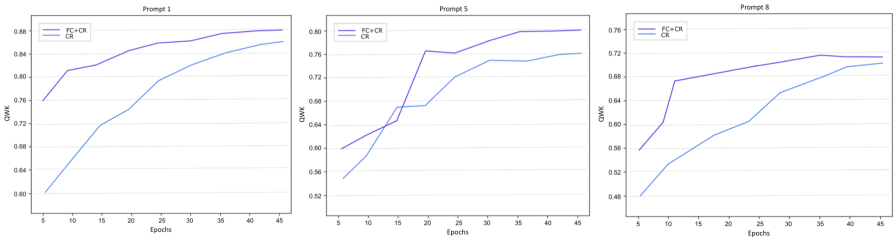


Fig. 2. Plot showing the increase in QWK metric performance with the increase in number of epochs on validation data during the training of prompt 1, 5 and 8 respectively.

The QWK score vs epoch plots in Figure 2, shows that even at the beginning of the training process, CR + FC has significantly faster and greater convergence than CR on the validation data. Furthermore, as the number of epochs increases, there is an increase in the QWK scores. Introducing handcrafted features to CapsRater has penalized responses based on rubrics and has made a definite difference in the scoring. Interestingly, the average score of FeatureCapture (FC) by itself is 7% lower than the CR+FC model. This signifies that FeatureCapture, independently, does not have an excellent performance. Table 4 and Figure 2 throw light on the importance of

Table 5. For prompt 1 question: *Write a letter to your local newspaper in which you state your opinion on the effects computers have on people*, the Table lists the original essay response along with adversarially perturbed reponses. Green shows the addition of colored lines to original response, while red shows deletion of the colored lines from the response.

		Example
Original Response		Dear Local Newspaper, I believe that computers are an extremely useful tool in society... Also, it lets you communicate with friends and family through the internet, for example, using facebook ... Each class teaches us to respect more and more of this culture... Also, is another website that helps us study for vocab. It also has many games that and learn our vocab words... I assume you will understand how much the computer has made a positive effect in society.
Category	Attack	
ADD	AddSongs	Dear Local Newspaper, I believe that computers are an extremely useful tool in society. It helps people learn new things about different cultures . . . So shine bright, tonight you and I, We're beautiful like diamonds in the sky. . . I assume you will understand how much the computer has made a positive effect in society.
ADD	Repeat-Sent	Dear Local Newspaper, I believe that computers are an extremely useful tool in society... Also, it lets you communicate with friends and family through the internet, for example, using facebook. Finally, it provides an accurate research tool for school projects, or interviews.Also, it lets you communicate with friends and family through the internet, for example, using facebook. Finally, it provides ...
DEL	DelRand	Dear Local Newspaper, I believe that computers are an extremely useful tool... It helps people learn new things about different cultures. Also, it lets you communicate with friends and family through the internet for example, using facebook... I assume you will understand how much the computer has made a positive effect in society.
MOD	Mod-Grammar	Dear Local Newspaper, They believe this computer are the extremely useful tool is society. They helps this people should had learns ... Also, They lets u communicate with friend fam through the internets, 4 examples, using facebook. Finally, they provides the accurates researcher tools 4 schools project, or interview...
GEN	Babel-Gen	Computer with abandonment has not, and in all likelihood never will be boisterous, irreverent, and arrogant. Why is paper so accumulated to pondering? The reply to this query is that electronic computer is eternally and hastily incensed...

CapsRater, and the significance of its routing mechanism in modelling the hierarchical textual relationships. Moreover, using BERT embedding for initial text representation helps it grasp the contextual understanding of sentences in the response. Overall, this exhibits that the proposed work gains tremendously from utilizing both CapsNet architecture and extracted features.

5.1 Testing with Adversarial Essays

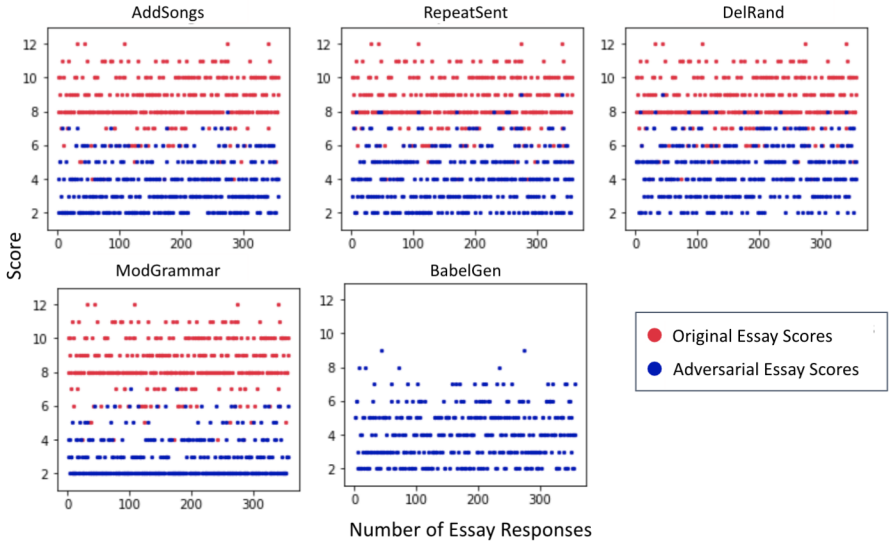


Fig. 3. CapsRater + FeatureCapture's scoring trends on Original and Adversarial essay responses for prompt 1

Recently, there has been thought-provoking work on the robustness of AES engines by conducting adversarial attacks on them [10, 19]. Kumar et al. [19] created a pipeline to generate various attacks on the ASAP-AES dataset. We employ their adversarially generated data to test the performance of our model. These are broadly divided into four categories: ADD (Text Addition), MOD (Text Modification), DEL (Text Deletion) and GEN (Random Text Generation). We experimented with two attacks from ADD and one from all other categories to check the robustness of the proposed model (CR + FC) and described the attacks with an example from prompt 1 response in Table 5. Here, *AddSongs* refers to the addition of lines of songs in the middle of the essay response. *RepeatSent* focused on repeating lines within the response essay to make it unreasonably wordy. *DelRand* created incoherent essays by removing random lines from the original response. *ModGrammar* introduced various grammatical errors in the original essay responses. *BabelGen* used a tool called the Babel Generator[7] to generate random sentences using prompt-specific keywords. *AddSongs*, *RepeatSent* and *DelRand* changed 25% of the original responses, while *ModGrammar* and *BabelGen* changed the entire original response. The training was done

[7] https://babel-generator.herokuapp.com/

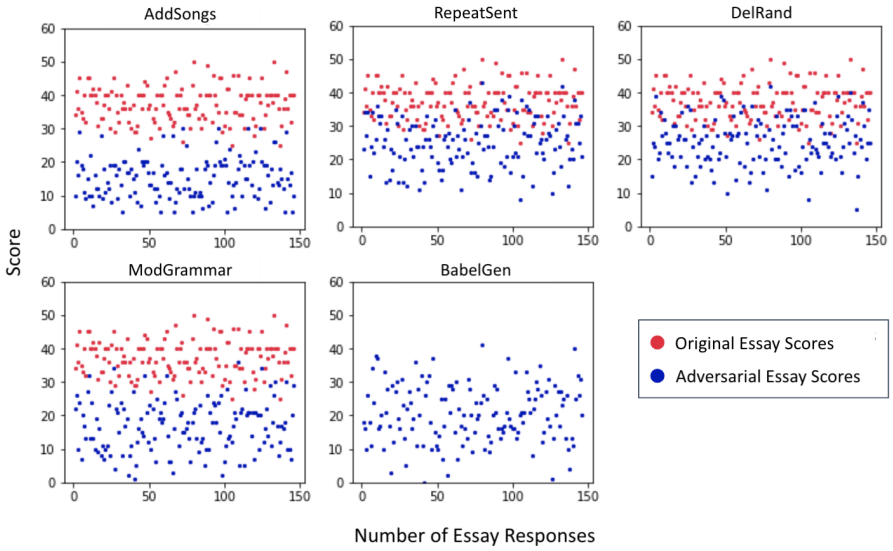


Fig. 4. CapsRater + FeatureCapture's scoring trends on Original and Adversarial essay responses for prompt 8.

on the original training data, while the testing data was replaced with the perturbed testing data according to the five selected adversarial attacks. To keep the analysis concise, we have shown the difference between the grading of Prompt 1 and 8's original responses and their corresponding adversarially perturbed responses. The results are shown in Figure 3 and 4.

Analysis: Original responses for prompt 1 are scored between the 2-12 range, and prompt 8 are scored between 0-60. We see that 90% of the original samples are scored between 8 and 12 for prompt 1. On similar lines, 88% of the original samples are scored between 30 and 60 for prompt 8. The overall trend is that adversarially perturbed responses have been scored consistently lower by CR+FC for all the attacks than the original essay responses. However, the maximum penalty was observed for the *ModGrammar* attack where 53% of responses of prompt 1 was scored the lowest possible score (2) and more than 70% of responses of prompt 8 was scored less than 30 (out of 60). *ModGrammar* made the responses grammatically incorrect and deformed the sentence structure such that it became semantically incorrect. For example, Table 5 shows a sentence where *ModGrammar* has introduced a singular demonstrative before plural noun: *"this people"*, a verb after a modal verb: *"should had"*, changed the semantics by introducing *"2"*, *"y"*, *"their"* incorrectly in the example. Hence, these responses were strongly penalized by CapsRater + FeatureCapture, due to excessive grammatical errors and semantic inconsistencies, leading to low scores. Another interesting observation is that while *RepeatSent* creates repetition of lines within the response, it still generates a relevant and coherent response. Even then, the attack was successfully identified, and most of the perturbed responses were scored lower for both prompt 1 and prompt 8. 45% of prompt 8 responses were scored between 20-30. We attribute this success to similarity-based features. *AddSongs* adds irrelevant lines to the responses, which tampers their relevance, structure and organization. We can see a clear drop in scores. Close to 65% of the perturbed responses for prompt 1 are scored between 2 and 4. Moreover, 40% of the responses for prompt 8 is scored between 15-20. This shows that the proposed work has grasped structural

inconsistencies and penalized such responses. Similarly, for *DelRand*, most of the perturbed responses are scored lower. We noticed that the prompt 1 score bracket is inclined towards the range of 4-6, and that of prompt 8 is between 20-30. Deleting random lines from an essay disturbs the structure and transition within sentences. While these perturbed essays may lack feature-based flaws, they were identified and penalized by CapsRater. Most interestingly, *BabelGen* attack was also captured by CR + FC. Babel Generator generates semantically correct but random passages using three keywords taken from the prompt. This type of attack does not have any spelling or word level inconsistencies. However, the inter-sentence coherence is extremely low, as shown in Table 5, by comparing the text in the original response and the *BabelGen* output. We observe that 80% of the responses' in this attack were scored lower (between 2-5) than original data for prompt 1. Similarly, 85% of prompt 8 responses were scored lower (between 5 and 30).

6 Conclusion

This work demonstrates an end-to-end pipeline that applies BERT enhanced Capsule Neural Network and handcrafted features for Automatic Essay Scoring. The CapsRater + FeatureCapture model reports a significant increase in the state-of-the-art performance on the QWK metric. Moreover, we conducted deeper experimentation and analyzed our technique's robustness against several types of adversarial attacks. CapsRater + FeatureCapture can detect and score the adversarial essay responses low. We aim to study and implement a domain-independent scoring system to eliminate training costs for each new question statement while retaining the performance in future work.

References

1. Dimitrios Alikaniotis, Helen Yannakoudakis, and Marek Rei. "Automatic text scoring using neural networks". In: *arXiv preprint arXiv:1606.04289* (2016).
2. Yigal Attali. *Automated Essay Scoring With E-rater® v.2.0*. URL: https://www.ets.org/Media/Research/pdf/RR-04-45.pdf.
3. Yigal Attali and Jill Burstein. "Automated essay scoring with e-rater® V. 2". In: *The Journal of Technology, Learning and Assessment* 4.3 (2006).
4. Barrust. *Pure Python Spell Checking*. URL: https://github.com/barrust/pyspellchecker.
5. Hongbo Chen and Ben He. "Automated essay scoring by maximizing human-machine agreement". In: *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*. 2013, pp. 1741–1752.
6. Hongbo Chen, Jungang Xu, and Ben He. "Automated essay scoring by capturing relative writing quality". In: *The Computer Journal* 57.9 (2014), pp. 1318–1330.
7. Tianqi Chen and Carlos Guestrin. "Xgboost: A scalable tree boosting system". In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. 2016, pp. 785–794.
8. Mădălina Cozma, Andrei M Butnaru, and Radu Tudor Ionescu. "Automated essay scoring with string kernels and word embeddings". In: *arXiv preprint arXiv:1804.07954* (2018).
9. Jacob Devlin et al. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. 2019. arXiv: 1810.04805 [cs.CL].

10. Yuning Ding et al. "Don't take "nswvtnvakgxpm" for an answer–The surprising vulnerability of automatic content scoring systems to adversarial input". In: *Proceedings of the 28th International Conference on Computational Linguistics*. 2020, pp. 882–892.
11. Edx EASE. *EASE (Enhanced AI Scoring Engine) is a library that allows for machine learning based classification of textual content. This is useful for tasks such as scoring student essays.* URL: https://github.com/edx/ease.
12. Youmna Farag, Helen Yannakoudakis, and Ted Briscoe. "Neural automated essay scoring and coherence modeling for adversarially crafted input". In: *arXiv preprint arXiv:1804.06898* (2018).
13. Geoffrey E Hinton, Alex Krizhevsky, and Sida D Wang. "Transforming auto-encoders". In: *International conference on artificial neural networks*. Springer. 2011, pp. 44–51.
14. Mohamed Abdellatif Hussein, Hesham Hassan, and Mohammad Nassef. "Automated language essay scoring systems: a literature review". In: *PeerJ Computer Science* 5 (2019), e208.
15. Robin Jia and Percy Liang. "Adversarial examples for evaluating reading comprehension systems". In: *arXiv preprint arXiv:1707.07328* (2017).
16. Cancan Jin et al. "TDNN: a two-stage deep neural network for prompt-independent automated essay scoring". In: *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 2018, pp. 1088–1097.
17. Zixuan Ke and Vincent Ng. "Automated Essay Scoring: A Survey of the State of the Art." In: *IJCAI*. 2019, pp. 6300–6308.
18. Jaeyoung Kim et al. "Text classification using capsules". In: *Neurocomputing* 376 (2020), pp. 214–221.
19. Yaman Kumar et al. "Calling out bluff: Attacking the robustness of automatic scoring systems with simple adversarial testing". In: *arXiv preprint arXiv:2007.06796* (2020).
20. Leah S Larkey. "Automatic essay grading using text categorization techniques". In: *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*. 1998, pp. 90–95.
21. Jiawei Liu, Yang Xu, and Yaguang Zhu. "Automated essay scoring based on two-stage learning". In: *arXiv preprint arXiv:1901.07744* (2019).
22. Maxbachmann. *Rapid Fuzz Similarity Calculator*. URL: https://github.com/maxbachmann/rapidfuzz.
23. Tomas Mikolov et al. "Distributed representations of words and phrases and their compositionality". In: *arXiv preprint arXiv:1310.4546* (2013).
24. Sara Sabour, Nicholas Frosst, and Geoffrey E Hinton. "Dynamic routing between capsules". In: *arXiv preprint arXiv:1710.09829* (2017).
25. Tulika Saha et al. "BERT-Caps: A Transformer-Based Capsule Network for Tweet Act Classification". In: *IEEE Transactions on Computational Social Systems* 7.5 (2020), pp. 1168–1179.
26. Swapna Somasundaran, Jill Burstein, and Martin Chodorow. "Lexical chaining for measuring discourse coherence quality in test-taker essays". In: *Proceedings*

of COLING 2014, the 25th International conference on computational linguistics: Technical papers. 2014, pp. 950–961.
27. Kaveh Taghipour and Hwee Tou Ng. "A neural approach to automated essay scoring". In: *Proceedings of the 2016 conference on empirical methods in natural language processing*. 2016, pp. 1882–1891.
28. Yi Tay et al. "Skipflow: Incorporating neural coherence features for end-to-end automatic text scoring". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 32. 1. 2018.
29. Viraj. *Python wrapper for Grammar Checking*. URL: https://pypi.org/project/grammar-check/1.3.1/.
30. Yucheng Wang et al. "Automatic essay scoring incorporating rating schema via reinforcement learning". In: *Proceedings of the 2018 conference on empirical methods in natural language processing*. 2018, pp. 791–797.
31. Ruosong Yang et al. "Enhancing Automated Essay Scoring Performance via Cohesion Measurement and Combination of Regression and Ranking". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Findings*. 2020, pp. 1560–1569.
32. Helen Yannakoudakis, Ted Briscoe, and Ben Medlock. "A new dataset and method for automatically grading ESOL texts". In: *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies*. 2011, pp. 180–189.
33. Torsten Zesch, Michael Wojatzki, and Dirk Scholten-Akoun. "Task-independent features for automated essay grading". In: *Proceedings of the tenth workshop on innovative use of NLP for building educational applications*. 2015, pp. 224–232.
34. Siyuan Zhao et al. "A memory-augmented neural model for automated grading". In: *Proceedings of the fourth (2017) ACM conference on learning@ scale*. 2017, pp. 189–192.
35. Wei Zhao et al. "Investigating capsule networks with dynamic routing for text classification". In: *arXiv preprint arXiv:1804.00538* (2018).
36. Wei Zhao et al. "Towards scalable and reliable capsule networks for challenging NLP applications". In: *arXiv preprint arXiv:1906.02829* (2019).

Journal of Chemical Neuroanatomy 116 (2021) 102012

Contents lists available at ScienceDirect

Journal of Chemical Neuroanatomy

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

FOXO and related transcription factors binding elements in the regulation of neurodegenerative disorders

Vaibhav Oli [1], Rohan Gupta [1], Pravir Kumar *

Molecular Neuroscience and Functional Genomics Laboratory, Delhi Technological University (Formerly Delhi College of Engineering), India

ARTICLE INFO

Keywords:
Alzheimer's disease
Parkinson's disease
FOXO
Oxidative stress
Insulin signaling
PI3K/AKT pathway
DAF-16
Natural biomolecules

ABSTRACT

Neurodegenerative diseases, such as Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and others, are characterized by progressive loss of neuronal cells, which causes memory impairment and cognitive decline. Mounting evidence demonstrated the possible implications of diverse biological processes, namely oxidative stress, mitochondrial dysfunction, aberrant cell cycle re-entry, post-translational modifications, protein aggregation, impaired proteasome dysfunction, autophagy, and many others that cause neuronal cell death. The condition worsens as there is no effective treatment for such diseases due to their complex pathogenesis and mechanism. Mounting evidence demonstrated the role of regulatory transcription factors, such as NFκβ, FoxO, Myc, CREB, and others that regulate the biological processes and diminish the disease progression and pathogenesis. Studies demonstrated that forkhead box O (FoxO) transcription factors had been implicated in the regulation of aging and longevity. Further, the functions of FoxO proteins are regulated by different post-translational modifications (PTMs), namely acetylation, and ubiquitination. Various studies concluded that FoxO proteins exert both neuroprotective and neurotoxic properties depending on their regulation mechanism and activity in the brain. Thus, understanding the nature of FoxO expression and activity in the brain will help develop effective therapeutic strategies. Herein, firstly, we discuss the role of FoxO protein in cell cycle regulation and cell proliferation, followed by the regulation of FoxO proteins through acetylation and ubiquitination. We also briefly explain the activity and expression pattern of FoxO proteins in the neuronal cells and explain the mechanism through which FoxO proteins are rescued from oxidative stress-induced neurotoxicity. Later on, we present a detailed view of the implication of FoxO proteins in neurodegenerative disease and FoxO proteins as an effective therapeutic target.

Abbreviations: NDDs, Neurodegenerative disease; AD, Alzheimer's disease; PD, Parkinson's disease; ALS, Amyotrophic lateral sclerosis; HD, Huntington's disease; FoxO, Forkhead box protein; CREB, cAMP-response element-binding protein; NFκβ, Nuclear factor kappa-light-chain-enhancer of activated B cells; NRF2, Nuclear factor erythroid 2-related factor 2; Aβ, β-amyloid; GDNF, Glial derived neurotrophic factor; MST1, Macrophage stimulating 1; JNK, c-Jun N-terminal kinase; AMPK, 5′ AMP-activated protein kinase; IGF, Insulin-like growth factor; PTMs, Post-translational modifications; USP1, Ubiquitin specific peptidase 1; EPCs, Endothelial progenitor cells; CDK, Cyclin-dependent kinases; FACS, Fluorescence-activated cell sorting; PCNA, Proliferating cell nuclear antigen; HAT's, Histone acetyltransferase; HDACs, Histone deacetylase; SOD2, Superoxide dismutase; Skp2, S-phase kinase-associated protein 2; CHIP, C-terminus of Hsc70-interacting protein; MDM2, Murine double minute 2; COP1, Constitutive photomorphogenic 1; TRAF3, TNF receptor-associated factor 3; TBK1, TANK binding kinase 1; IRF3, IFN regulatory factor 3; TREM2, Triggering receptor expressed on myeloid cells 2; ROS, Reactive oxygen species; NADPH, Nicotinamide adenine dinucleotide phosphate; Nox, NADPH oxidase; XO, Xanthine oxidase; ER, Endoplasmic reticulum; LPS, Lipopolysaccharides; APP, Amyloid precursor protein; PS1, Presenilin 1; SNpc, Substantia nigra pars compacta; PARP, Poly (ADP-ribose) polymerase; NLS, Nuclear localization signal; PUMA, p53 upregulated modulator of apoptosis; DG, Dentate gyrus; PRMT1, Protein arginine N-methyltransferase 1; PGC-1α, Peroxisome proliferator-activated receptor-gamma coactivator 1-alpha; XBP1, X-box-binding protein 1; FBXO32, F-Box Protein 32.

* Corresponding author at: Professor and Head, Department of Biotechnology, Delhi Technological University (Formerly Delhi College of Engineering); Molecular Neuroscience and Functional Genomics Laboratory, Shahbad Daulatpur, Bawana Road, Delhi, 110042, India.
E-mail address: pravirkumar@dtu.ac.in (P. Kumar).

[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jchemneu.2021.102012
Received 16 May 2021; Received in revised form 16 July 2021; Accepted 7 August 2021
Available online 13 August 2021
0891-0618/© 2021 Elsevier B.V. All rights reserved.

Table 1
Some essential mouse models for Alzheimer's disease (AD), Parkinson's disease (PD), and Huntington's disease (HD).

Mouse models	Disease	Mechanism	Type	References
rTg4510 mice	AD	A reversible binary transactivator system produces tau overexpression.	Genetic (Random transgene integration)	(SantaCruz, 2005)
KO based mice	AD	Open reading frames of genes like APP, MAPT, BACE1 are disrupted.	Genetic (HDR-based)	(Zheng et al., 1995)(Harada et al., 1994)(Cai et al., 2001)
3×Tg strain of mice	AD	Two constructs are injected simultaneously, and one is APP expressing, and the other is P301 l mutant tau type.	Combinatorial genetic	(Oddo et al., 2003)
hTau mice (Trauma-based)	AD	Injury or trauma to the brain for tau pathology development.	Non- genetic	(Ojo et al., 2013)
LRRK2 mutations-based mice	PD	Inclusion formation is affected and slight dopaminergic neurodegeneration.	Genetic (Viral vector-mediated)	(Blesa and Przedborski, 2014)(Dawson et al., 2010)
SNCA transgenic mice	PD	A large amount of α-synuclein aggregation.	Genetic (Point mutation)	(Fleming et al., 2005)
UCH-L1 mutations-based mice	PD	Dopaminergic neurodegeneration induced.	Genetic (I93 M mutation)	(Setsuie et al., 2007)
Nurr1 deficient mice	PD	Dopaminergic neurodegeneration is induced.	Genetic (Gene knock-out)	(Jiang et al., 2005)
R6/2 mice	HD	N-terminal fragment of htt gene (exon-1) with 144−150 CAG repeats at exon-1 is expressed.	Genetic (transgenic)	(Mangiarini et al., 1996)
N171−82Q mice	HD	N-terminal fragments with exon-1 and exon-2 of the htt gene with 82 polyglutamine are expressed.	Genetic (transgenic)	(Schilling, 1999)
HdhQIII	HD	111 CAG repeats are inserted into the murine HD gene.	Genetic (Gene knock-in)	(Wheeler, 2000)
BACHD mice	HD	170 kb of human huntingtin locus is expressed.	Genetic (transgenic)	(Gray et al., 2008)

1. Introduction

Life expectancy has increased drastically owing to the advances in medical sciences. Still, with this, the old age-dependent diseases, namely neurodegenerative disease (NDDs), have started to become predominant (Heemels, 2016). NDDs, namely Alzheimer's disease (AD), Parkinson's disease (PD), Huntington's disease (HD), and amyotrophic lateral sclerosis (ALS), frontotemporal dementia (FTD), are caused by excessive loss of neurons, which leads to memory impairment and cognitive dysfunction (Peng et al., 2020). However, due to their complex etiology, researchers across the globe remain failed to find an effective treatment for NDDs (Gitler et al., 2017). Different experimental model organisms have been designed to understand the genomics and proteomics of the disease pathogenesis from the past several decades and develop effective therapeutic strategies. Table 1 summarizes a few essential genetic and non-genetic mouse models of NDDs.

Recently the application of transcription factors, namely forkhead box (FoxO), cAMP-response element-binding protein (CREB), nuclear factor kappa-light-chain-enhancer of activated B cells (NFκβ), and many others as effective therapeutic targets has been demonstrated in multiple NDDs models. For instance, activation of NFκβ cascade through upregulation of fibroblast growth factor 9 promotes neuronal development and neurite outgrowth in HD model, whereas, downregulation of NFκβ through upregulation of nuclear factor erythroid 2-related factor 2 (NRF2) and heme oxygenase 1 prevents from β-amyloid (Aβ)-induced neurotoxicity (Ali et al., 2021)(Yusuf et al., 2021). Similarly, activation of CREB signaling causes upregulation of glial-derived neurotrophic factor (GDNF) and dopamine in the PD cell culture model, whereas, activation of CREB signaling rescued from Aβ pathology and related cognitive decline through enhanced autophagic degradation (Katila et al., 2020)(Wang et al., 2021). Further, from the past decade involvement of FoxO transcription factors as a therapeutic target in the pathogenesis of NDDs has been extensively studied. For example, depletion of *drosophila* CREB-binding protein dCBP protects glial cell through FoxO-dependent signaling event in HD pathology, whereas, upregulation of FoxO decreases aging and prevent axonal degeneration through enhanced autophagy (Martin et al., 2021)(Hwang et al., 2018). FoxO regulates diverse physiological and pathological mechanisms involved in age-dependent degeneration, namely oxidative stress, cell cycle entry, autophagy, apoptosis, angiogenesis, and mitochondrial homeostasis (Tia et al., 2018).

FoxO family are a set of transcription factors that have a winged helix-like structure for its DNA binding domain named forkhead box, and the FoxO is a subclass of the complete FOX family (Kaestner et al., 2000). FoxO family in mammals include FoxO1, FoxO3, FoxO4, FoxO6 but other organisms like drosophila contain only one member. They are different from each other only in their expression within different cellular and tissue-specific environments. So basically, they are transcriptional activators, and they are regulated by signaling cascades of insulin and growth factors. FoxO transcription factors are activated by the different signaling events, such as PI3K/Akt pathway, macrophage stimulating 1 (MST-1) stress response pathway, c-Jun N-terminal kinase (JNK) pathway, Sirtuin 2 cascade, and 5′ AMP-activated protein kinase (AMPK) signaling pathway under different physiological conditions (Neri, 2012). The targets of FOXO, which were discovered first, include genes for metabolism and stress resistance as whenever the cells are not generating enough energy or when insulin or insulin-like growth factors (IGF) are not present. FoxO accumulates in the nucleus and stimulates activation of such genes (Kops et al., 1999)(Tran, 2002)(Nemoto, 2002). Further, studies demonstrated that FoxO shuttles between the nucleus and cytoplasm. For instance, phosphorylation of FoxO through Akt causes its translocation to the nucleus. In addition, the activity of FoxO is regulated by several post-translational modifications (PTMs), namely acetylation, methylation, deacetylation, ubiquitination, phosphorylation, glycosylation, and SUMOylation. For example, CBP-induced FoxO acetylation inhibits FoxO ubiquitination and protects from oxidative stress, whereas, sirtuin 1-induced activation of FoxO inhibits oxidative stress and apoptosis (Qin et al., 2021)(Zou et al., 2021). Similarly, under prolonged starvation, ubiquitin specific peptidase 1 (USP1) inhibits polyubiquitination of Akt at K63, which inhibits PI3K/Akt/FoxO signaling events (Goldbraikh et al., 2020). However, apart from phosphorylation, the role of PTMs in the regulation of FoxO activity is still devoid. Thus, more *in* vivo and *in vitro* studies will be required to exploit the involvement of PTMs. Currently, attempts have been made to exploit the role of FoxO protein as a potential therapeutic target and several compounds, namely ligustilide, monascin, trehalose, metformin have been demonstrated as a potential activator of FoxO in the treatment of NDDs. Fig. 1 gives a schematic representation of etiopathology, risk factors and development of NDDs with key therapeutic approaches for their treatment.

2. FOXO factors in cell proliferation, cell cycle, and regulatory proteins

In coronary heart disease, it is known that the destruction of vascular endothelial cells is an essential vascular pathophysiological aspect. Endothelial progenitor cells (EPCs) that circulate in the blood can move towards the injured site of ischemic tissue or blood vessels and keep the

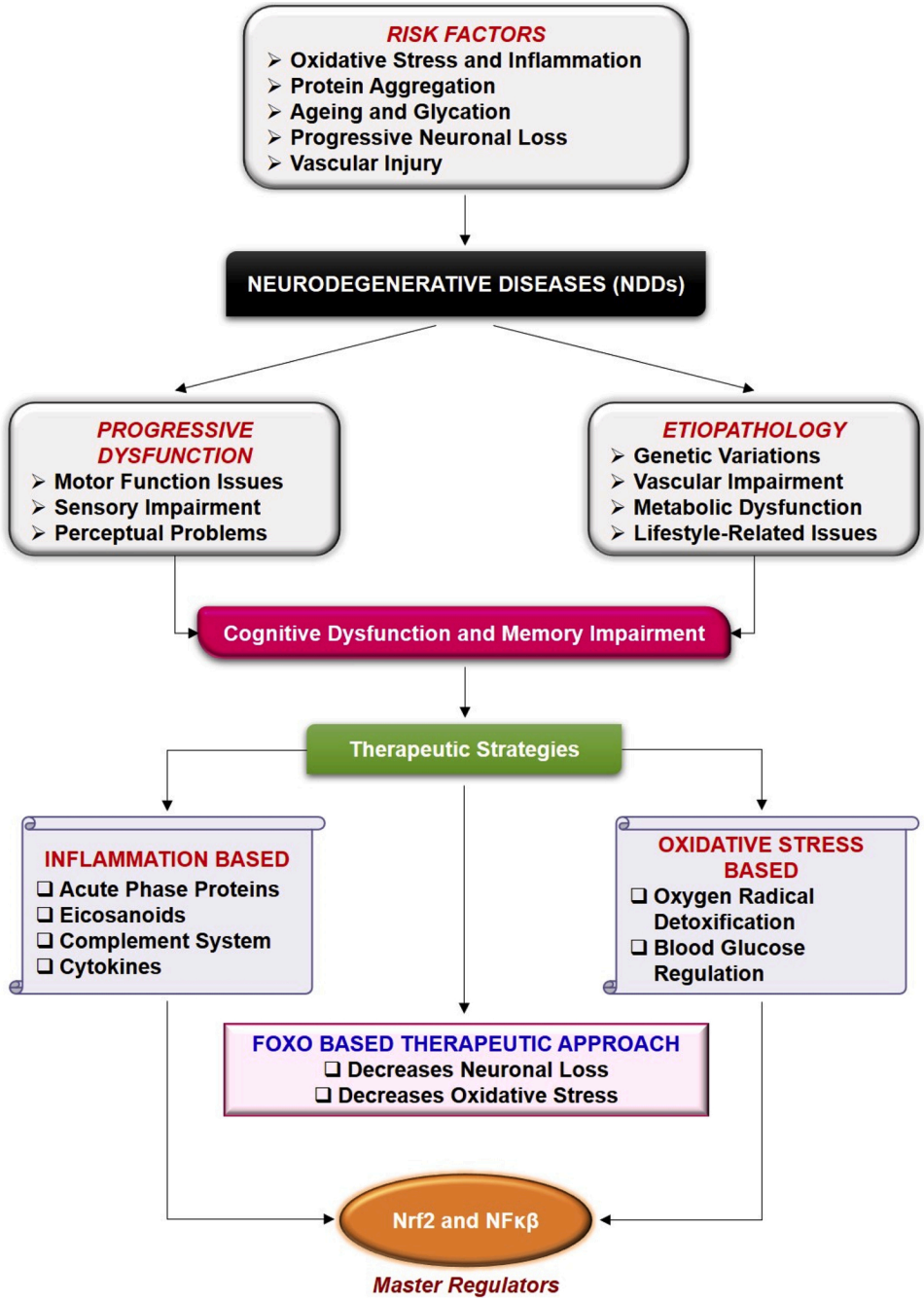


Fig. 1. Schematic diagram of the risk factors, etiopathology, causes, and critical therapeutic approaches for NDDs. Aging, oxidative stress, and inflammation, among others, make an individual prone to NDDs. Also, particular lifestyle adopted, genetic factors, vascular and metabolic factors can induce neurodegenerative diseases. The brain's motor, sensory, and perceptual function is affected by such disorders, leading to cognitive and behavioral decrement, a significant symptom of these disorders. Therapeutic approaches can be designed based on reducing inflammation and oxidative stress in the neuronal cells. Fundamental approaches can be cytokine therapy for inflammation and oxygen radical detoxification for oxidative stress. NRF2 and NF-κB can regulate both inflammation and oxidative stress.

integrity of endothelial cells, thereby differentiating into mature ones. This can be seen as a regenerative medicine approach (Urbich and Dimmeler, 2004). But patients of this disease usually have impaired EPCs. Their number is also low in circulation (Werner et al., 2005) due to aging and other cardiovascular risk factors (Tousoulis et al., 2008). So, by therapeutic intervention, we need to increase the number of EPCs and make them functional. Many studies have demonstrated in glioma cells, endothelial cells, and vascular smooth muscle cells that FoxO factors can negatively regulate proliferation (Lau et al., 2009). Also, it was found that FoxO3a expression is increased during oxidative stress in the EPCs, but there is no effect on FoxO1 and FoxO4 (Wang et al., 2013). FoxO3a is activated upon dephosphorylation by Akt and inhibits cell proliferation as it enters into the nucleus from the cytoplasm. Cyclin-dependent kinase (CDK) inhibitor p27kip1 at the protein level is accumulated due to which cyclin/CDK complexes are prevented, which are required to progress into the S phase by the cell. As a result, the cell cycle is arrested at Go/G1 phase (Wang et al., 2013). In a study by Tiantian Sang, Qing Cao et al., they have explored the mechanism of the proliferation of EPCs and regulation of cell cycle proteins by overexpression and inhibition of FOXO3a, respectively. In this study, EPCs were first isolated from the umbilical cord of a human donor and were then cultured in vitro. Three recombinant adenovirus vectors were made, namely Ad-TM (triple mutant)-FOXO3a, Ad-shRNA-FOXO3a, and Ad-GFP. In the Ad-TM (triple mutant)-FOXO3a, Thr-32, Ser-253, and Ser-315 are the sites of Akt phosphorylation replaced by alanine residues so that Akt cannot phosphorylate this type of construct. After that transfection was done, then the transfected cells were stored for further investigation. Flow cytometry was performed using propidium iodide stain for cell cycle assay, and fluorescence-activated cell sorting (FACS) was used to measure DNA content. Finally the western blot analysis was

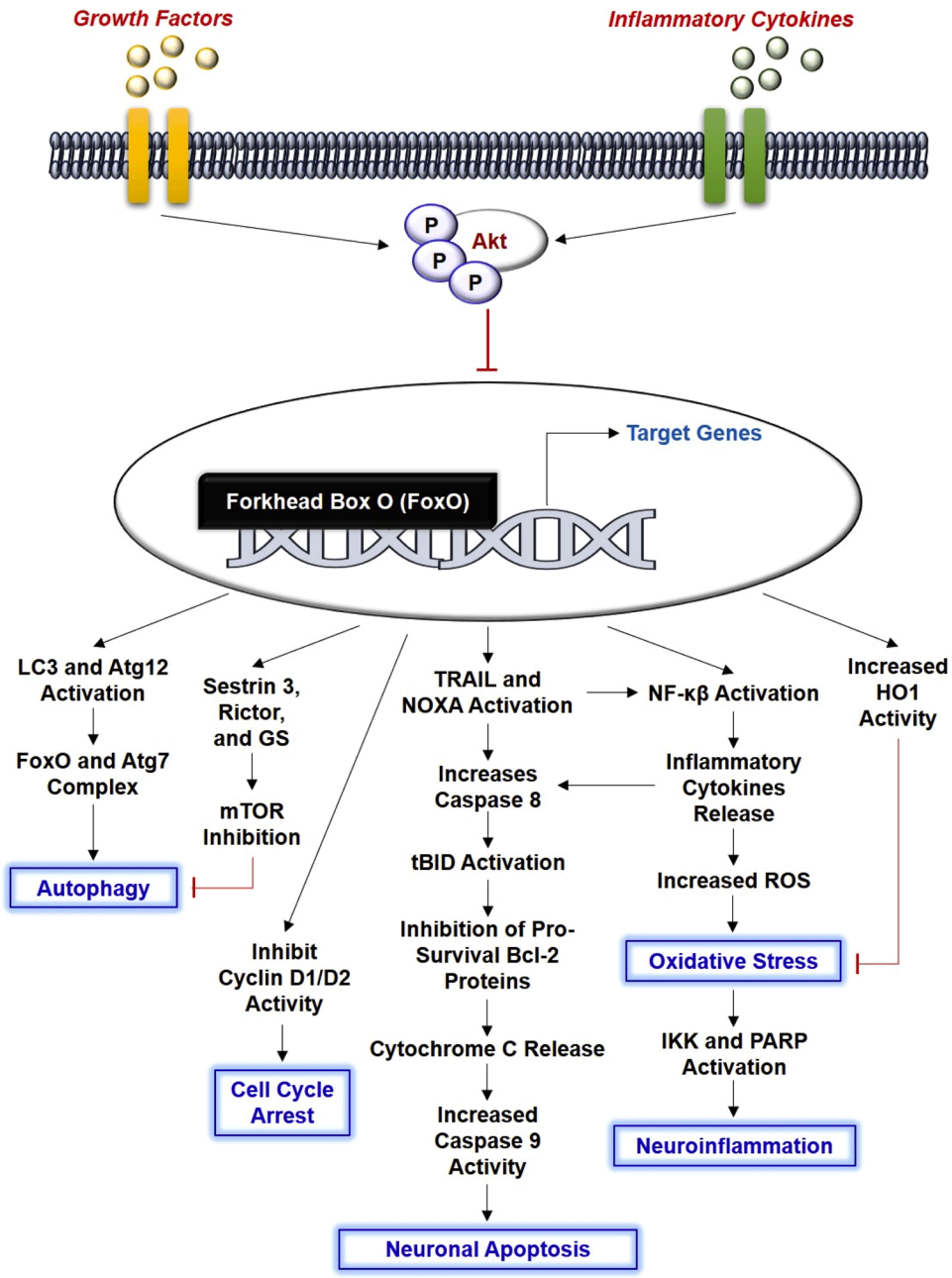


Fig. 2. FOXO regulates major cellular processes by regulating the expression of downstream genes. Various protein kinases and signaling proteins interact with FOXO in the cell's cytoplasm to activate or inhibit it for nuclear translocation. STAT3, p38, among others, can activate FOXO in the cytoplasm, whereas Akt, SET9, NLK, ERK1/2 can inhibit activation of FOXO. After the translocation of FOXO in the nucleus, it interacts with downstream genes. It regulates them to control major cellular processes like metabolism, apoptosis, and various others for maintaining the necessary functions and life of the cell.

performed to obtain results (Sang et al., 2014).

Transient transfection of CCK-8 assay was done to check the effect of FoxO3a on EPC proliferation, and for that cell, viability readings were taken after 24, 48, and 72 h. At the three time points, transfection with Ad-TM-FOXO3a resulted in a substantial reduction in EPC proliferation compared to Ad-GFP transfection. Also, the results with transfection with Ad-shRNA-FoxO3a show substantial enhancement in the proliferation of EPCs. So it was concluded from the CCK-8 assay that EPC proliferation is reduced by the presence of FoxO3a (Ojo et al., 2013). Further, flow cytometry observations to check the effect of FoxO3a on the cell cycle show that Ad-TM-FoxO3a transfected EPCs clearly show arrest at the G1 phase. When FoxO3a is silenced, there is a significant increase in the proliferation of EPCs and the percentage of cells in the S and G2 phases. The proliferation index is very high compared to Ad-GFP transfected EPCs. This means the silencing of FoxO3a triggers the cell cycle progression. So this can be explained in this way that PI3K/Akt signaling pathway related to FoxO3a prevents cell cycle progression from G1 to S phase, and thus the proliferation of EPCs is blocked (Ojo et al., 2013).

After transfection, a western blot was performed to check the effects of FoxO3a on regulatory proteins of the cell cycle. The expression of proliferating cell nuclear antigen (PCNA) protein, which is related to the synthesis of DNA in the EPCs, significantly decreased in the Ad-TM group and significantly increased in the Ad-shRNA group compared to GFP, the control group . Thus, it was concluded that FOXO3a could reduce cell proliferation in EPCs and can be related to PCNA protein levels. In addition to this, it was found that p27kip1, which regulates cell cycle transition from G1 to S phase (Toyoshima & Hunter, 1994) and is a cyclin-dependent kinase inhibitor, was upregulated by overexpression of FOXO3a and was downregulated when FOXO3a was inhibited. Thus, it can be concluded that in EPCs, the expression of p27kip1 protein is seen when there is overexpression of FOXO3a. CDs like p27kip1 suppress complexes like cyclin D1/CDK6 and cyclin E/CDK2 in G1 to S phase progression of the cell cycle (Chu et al., 2008). Western blots for cyclin

D1 and CDK2 expression show that EPCs transfected with Ad-TM-FoxO3a and Ad-shRNA-FoxO3a downregulate and upregulate their expression levels, respectively. So based on the presence or absence of FoxO3a, it may be possible that cell cycle regulatory proteins can be regulated by it. Thus, it can control the transition of the G1 to S phase, blocking the proliferation of EPCs (Sang et al., 2014). This may be applied to other body cells and neuronal cells as a regenerative medicinal approach. Various protein kinases and signaling molecules that activate and inhibit FoxO, which leads to the regulation of genes that control major cellular processes, are shown in Fig. 2.

3. Involvement of lysine post-translational modification in FoxO regulation

It is a well-established notion that the activity of FoxO transcriptional factors is regulated by PTMs, where lysine PTMs, such as acetylation, deacetylation, ubiquitination, methylation, and SUMOylation plays a crucial role in the regulation of FoxO transcriptional activity. Acetylation and deacetylation of FoxO proteins are carried out with the help of enzymes called histone acetyltransferase (HATs) and histone deacetylase (HDACs) (Daitoku et al., 2011). For instance, acetylation of FoxO3 (K242 and K245) and FoxO1 (K245 and K248) decreases the DNA binding activity of FoxO that contributes to the regulation of signaling pathways (Brown and Webb, 2018). Further, acetylation with CBP/p300 decreases the activity of FoxO but increases the transcription of genes through FoxO (Zhao et al., 2011). Additionally, acetylation through sirtuin 1 and sirtuin 2 regulates the signaling events involved in longevity (Horst et al., 2004). HDAC3 activates the FoxO and promotes the transcription of target genes through deacetylating activity. Recently, Chang et al., 2019 demonstrated that genetic removal of sirtuin 3 reversed the acetylation of FoxO3 and superoxide dismutase 2 (SOD2), which protects from oxidative stress and apoptosis (Chang et al., 2019). Ubiquitination is another lysine PTM, which carried out the elimination of FoxO proteins after phosphorylation and activation by Akt and ERK. For example, an E3 ubiquitin ligase, S-phase kinase-associated protein 2 (Skp2), and C-terminus of Hsc70-interacting protein (CHIP) colocalized with FoxO proteins in the nuclear and cytoplasm, respectively (Huang and Tindall, 2011). Similarly, murine double minute 2 (MDM2) and constitutive photomorphogenic 1 (COP1) interacted with FoxO proteins and colocalized between nucleus and cytoplasm under DNA damage and oxidative stress conditions (Brenkman et al., 2008)(Wang et al., 2016). A study concluded that phosphorylation of FoxO at S256 causes its ubiquitination through Skp2 (Wang et al., 2016). However, the ubiquitination mechanism of FoxO4 is slightly different from the ubiquitination mechanism of other FoxO proteins. A study by Wang et al., 2015 demonstrated that sirtuin 1 protects from oxidative stress-induced cell death by inhibiting the activity of FoxO3 by its ubiquitination (Wang et al., 2015). Similarly, in another study Wang et al., 2020a concluded that ISGylation but not ubiquitination of FoxO3a is increased in the presence of proteasome inhibitor (Wang et al., 2020a). In addition, Ma et al., 2021 reported that Inhibition of TNF receptor-associated factor 3 (TRAF3) ubiquitination at K63 through FoxO1 causes disassociation between TRAF3 and TANK binding kinase 1 (TBK1), which further destabilize IFN regulatory factor 3 (IRF3) (Ma et al., 2021). Thus, it can be concluded that PTM play a major role in the regulation of FoxO proteins activity both in nucleus and in cytoplasm.

4. Expression and regulatory functions of FoxO proteins in neuronal cells

Accumulated evidence suggests the critical role of FoxO proteins expression and activity in the neuronal cells, which regulates different physiological processes (Martins et al., 2016)(Wang et al., 2014b) (Urbánek and Klotz, 2017). Recent studies on *Caenorhabditis elegans* (*C. elegans)* have helped us to understand the expression and function of FoxO in neuronal cells. In these organisms, the FoxO ortholog, called DAF-16, is present, responsible for regeneration, stress resistance, memory, and learning in neurons. To investigate similar features of FoxO in neurons of mammals should be the next step for its better understanding. A therapeutic approach for NDDs can be designed if one knows the function of FoxO in neurons which are healthy, aged, or damaged by any one of such disorders (Kim and Webb, 2017). In neuroscience, *C. elegans* has emerged as a suitable model for FoxO studies because of its simplicity, neuronal wiring that is stereotyped, cell lineage invariance, and behavioral plasticity. Each of these worms has precisely 302 neurons, and consistent connectivity of cells between these animals can be observed (Tissenbaum, 2018). So based on its specific position, morphology, and connectivity, each neuron can be identified. Studies on the behavior of this worm and new approaches in genomics have identified the neuronal function of DAF-16/FoxO in memory and learning. For instance, studies demonstrated that DAF-16/FoxO activity is involved in neuronal compensation, cellular plasticity, and taste avoidance learning (Farina et al., 2017)(Nagashima et al., 2019)(Aghayeva et al., 2020). Further, DAF-16 in isothermal tracking experiments enhances associative learning, while in salt chemotaxis experiments, it has a slight effect to overcome the defect of daf-2. This suggests that other factors work downstream in insulin/IGF-1 signaling in the salt chemotaxis experiments.

Moreover, aging in *C. elegans,* like in humans, affects memory performance, but it was observed that daf-16 can regulate that. For instance, Kauffman et al., demonstrated that a positive olfactory assay was developed to assess the age-related effect on learning and memory in these animals. This chemo sensation is developed with AWS neurons which associate food with a particular odorant like butanone in the assay (Kauffman et al., 2010). The first neuronal function that is lost due to age is long olfactory memory, which is regulated by CREB and is similar in mammals. Further, Daf-2 mutants show 3 times more short-term memory and enhanced long-term memory when compared to wild types. Although such mutant's longevity can be observed, they do not deliver long-term memory performance. In addition, it is observed that insulin/IGF-1 signaling regulates longevity in intestine cells but not in neurons. This suggests that DAF-16 holds a particular set of specific genes to the neurons (Libina et al., 2003). More studies are ongoing to ascertain specific neurons that DAF-16 directly regulates. Furthermore, it will be essential to map the downstream mechanism of DAF-16 in different neuronal subtypes to understand why its expression is different for some neurons. They maintain longevity and young age-dependent expression (Kim and Webb, 2017).

The interesting question is that we can compare this DAF-16 expression and function with FoxO in mammals and humans. Although studies have not yet given any substantial evidence, experiments with knockout mice talk about its function in humans. Also, studies on cognitive processes have established a strong link between them and insulin/IGF signaling in humans. A specific allele of FoxO3 can be directly related to the increased cognitive ability of humans with age (Willcox et al., 2008). The four isoforms of FoxO, i.e., FoxO1, FoxO3, FoxO4, and FoxO6, are expressed (Hoekman et al., 2006). The mouse models to study the function of FoxOs in the brain have several results. For example, FoxO1 and FoxO3 have been found to regulate the anxiety-based behavior of the mouse (Polter et al., 2009). Further, Nucleus accumbens (ventral striatum), also called the brain's reward center, is also regulated by FoxO3, which was tested with behavioral response to cocaine (Ferguson et al., 2015). And finally, FoxO6 is limited to regulate learning and memory in mammals and humans (Salih et al., 2012).In addition, FoxO4 is the least expressed of all in the neurons of the brain (Biggs III and Cavenee Karen C., 2001). Isoforms of the FoxOs have similar DNA binding domains but very different in other regions. The FoxOs can interact with different cofactors to regulate the expression of the genes in the hippocampus in the brain and many cellular processes (Jacobs et al., 2003).

Our main aim is to investigate whether, like *C. elegans*, can FoxOs in mammals also can maintain neuronal function and protect them from

Table 2
Regulation of FoxO protein expression in the neuronal system and its corresponding therapeutic effect.

FOXO Protein	Neuronal System	Expression	Effect	References
FOXO1	Autophagy mediator	Upregulated	Promotes the autophagic degradation of toxic protein aggregates	(Chaubey et al., 2020)
	Neural stem/progenitor cells (NSPCs)	Downregulated	Promotes NSPCs differentiation into neurons	(Kim et al., 2015a)
	Oxidative stress response	Upregulated	Administration of pro-apoptotic protein apoptosis enhanced H_2O_2-induced cytotoxicity and promoted neuronal cell death through the JNK-FoxO1 pathway	(Zhang et al., 2019)
	Anti-apoptosis effects	Downregulated	Negative regulation of FoxO1 by miR181a protects Aβ-induced cell apoptosis in APP/PS1 mice	(Wu et al., 2019)
	Neuronal circuit reprogramming	Upregulated	Restore synaptic transmission and repair hippocampal cholinergic circuit in an AD mouse model	(Xia et al., 2019)
	Neuronal apoptosis	Upregulated	CDK5-induced phosphorylation at S247 promotes FoxO1 nuclear translocation, which activates pro-apoptotic genes	(Zhou et al., 2015)
	Neuronal reprogramming	Downregulated	Production of induced neuronal cells (iN) from aged mice fibroblasts	(Ahlenius et al., 2016)
FOXO3	Oxidative stress response	Upregulated	Promotes autophagic degradation in HT22 cell culture model to provide neuroprotection against H_2O_2-induced cytotoxicity	(Deng et al., 2020)
	Neurogenesis and neural differentiation	Upregulated	Inhibits cellular senescence through FoxO3/ETS2/p16 axis	(Voisin et al., 2020)
	Mitochondrial pathway	Downregulated	Genetic knockdown of cicFoxO3 protects the neuronal cell from oxidative stress through mitochondrial apoptosis signaling	(Lin et al., 2020)
	Neurogenesis	Upregulated	Promotes autophagic degradation and help in maintaining neurogenesis	(Audesse et al., 2019)
	Neuronal response	Upregulated	Activates neuronal signaling in response to AICD	(Bernard et al., 2018)
FOXO3a	Mitochondrial activity regulation (for all cells including neurons)	Upregulated	Limits ROS production from mitochondria during the hypoxic condition	(Ferber et al., 2012)
	Mitochondrial dysfunction	Upregulation	Induction of stress response mechanism and decreases oxidative stress	(Yang et al., 2021)
	Stress response mechanism	Upregulation	Activates stress response signaling in the presence of toxicity	(Reiterer and Milton, 2020)
	Stress response mechanism	Upregulation	Activates Sirt1/Sirt3/FoxO3a signaling and protects from neuronal toxicity	(Ruankham et al., 2019)
FOXO4	Oxidative stress response	Upregulated	FOXO4 and ATXN3 interaction with SOD2 increases antioxidative response	(Araujo et al., 2011)
	Regulate proteasome activity in neural circuit	Upregulated	Inhibiting Akt activity promotes FoxO4 and proteasome activity	(Liu et al., 2017)
	Neuronal morphogenesis	Downregulation	Causes impaired dendritic morphology and spine positioning	(Schäffner et al., 2018)
FOXO6	Memory consolidation	Upregulated	Regulation of gene expression for synaptic function and number control and also neuronal coordination in the hippocampus after memorizing	(Salih et al., 2012)
	Neurotransmission of pheromone signals	Upregulated	Maintains glutamate homeostasis in the hippocampus of mouse and head of *C. elegans*	(Park et al., 2017)
DAF-16/ FOXO	Oxidative stress response	Upregulated	Activates insulin signaling pathway and causes a decrease in stress conditions	(Hesp et al., 2015)
	Mitochondrial response	Upregulated	Improve health span and lifespan	(Yen et al., 2020)
	Lipid metabolism	Upregulation	Increases lipid metabolism	(Mutlu et al., 2020)
	Stress response mechanism	Upregulation	Expand lifespan through regulation of mitochondrial homeostasis	(Senchuk et al., 2018)
dFOXO	Oxidative stress sensitivity	Downregulated	DJ-1 inhibits dFOXO through PI3K/AKT, which leads to inhibition of DLP production and apoptosis in drosophila	(Hwang et al., 2013)
	Mitochondrial morphology and autophagic degradation	Upregulation	Promotes neuroprotection through maintenance of mitochondrial integrity	(Tas et al., 2018)
	Activates death signaling events	Downregulation	Inhibition of JNK/dFOXO pathway confer neuroprotection	(Hong et al., 2012)

the effects of aging. The role of FoxOs in aging and its regulation will be complex, but the expression profiling of the brain of humans has shown that there is dysregulation of synaptic genes due to increasing age (Lu et al., 2004). These genes are homologous to FoxO6 in humans, which is not confirmed, but FoxO6 does help in maintaining age-related effects of synaptic stability. Recently, Wang et al., 2020a,bdemonstrated the neuroprotective functions of triggering receptor expressed on myeloid cells 2 (TREM2) transcription factor. In the hippocampus of the 5xFAD AD mice model, activated the FoxO3a in the absence of TREM2 through PI3K/Akt pathway. Thus, the study concluded that neuroprotective functions of TREM2 are associated with the FoxO3a-PI3K axis (Wang et al., 2020b). Likewise, Bernard M et al., 2018 demonstrated that FoxO3a nuclear translocation is not directly proportional to AICD, but it is a direct consequence of Akt phosphorylation (Law et al., 2018). Apart from AD, the expression and activity of FoxO protein play an equally important role in the pathogenesis of HD. For instance, Kannike et al., 2014 demonstrated that 3-nitropropionic acid treatment in Hdh7/7 cells results in enhanced FoxO3a nuclear translocation, which regulates its transcription through binding with the FoxO3a promoter as a result of a positive feedback loop (Kannike et al., 2014). So, for now, it can be deduced from these studies that FoxO activation in the neurons can maintain cognitive function with aging. We can use their intentional stimulation as a therapeutic method to treat NDDs, but for that, more understanding of the underlying mechanisms is needed with extensive research. This will also enable us to more precisely define the different function of all the isoforms of FOXO in neurons in the different brain and spinal cord (Kim and Webb, 2017). Table 2 has an effort to summarize the expression of FoxO proteins in neuronal system and its corresponding effects that can help to design therapeutic approach for NDDs.

5. Oxidative stress and ROS in neuronal cells in neurodegenerative disorders

NDDs are characterized by the progressive loss of neurons with the

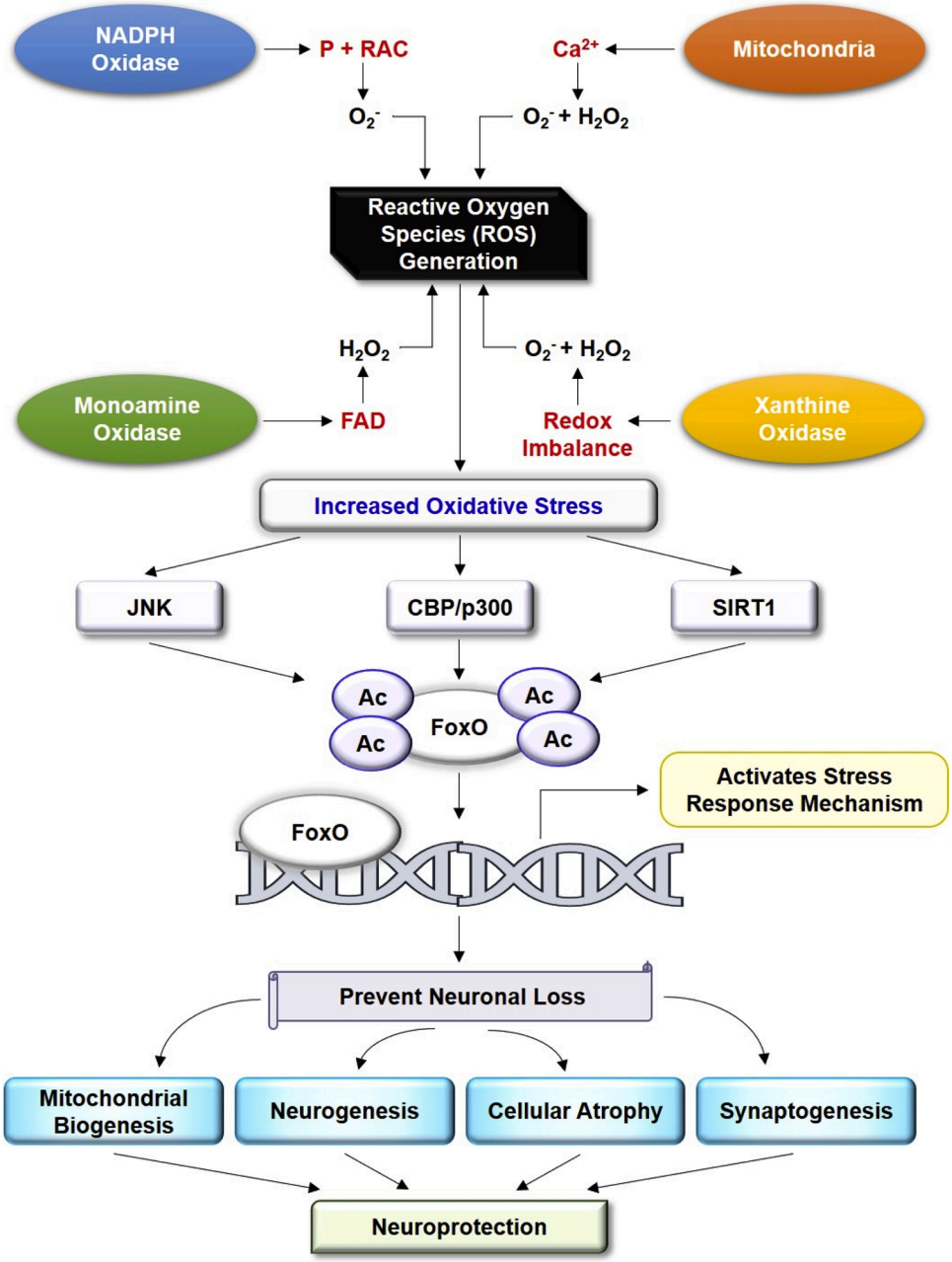


Fig. 3. Major sources of ROS production and their effects in the brain. NADPH oxidase, Xanthine oxidase, Monoamine oxidase, and mitochondria are major ROS producers for the brain and lead to the accumulation of superoxide (O_2-) and hydrogen peroxide (H_2O_2) in the brain. ROS accumulation in the brain results in oxidative stress in the neuronal cells. This results in an increase of 8-hydroxydeoxyguanosine, malondialdehyde, 4-hydroxynonenal, and protein carbonyls in the brain. ROS-mediated injury and oxidative protein damage in the brain lead to neurodegeneration.

increasing age, where oxidative stress is considered a significant contributor to disease pathogenesis (Gandhi and Abramov, 2012)(Lin and Beal, 2006). Oxidative stress can cause cell damage and malfunction the DNA repair mechanism of the cell, which leads to mitochondrial dysfunction (Song and Zou, 2015). The brain has plenty of peroxidation susceptible lipid cells and has excessive oxygen demand, so it is at risk for damage by reactive oxygen species (ROS). In NDDs, ROS-induced oxidative stress plays a significant role, and NDDs outcomes have been very inconsistent (Kim et al., 2015b). In the brain, both endogenous and exogenous sources are responsible for ROS production through chemical deposition, ionizing radiations, ultraviolet rays, and toxic pollutants (Mani, 2015). Mitochondria and enzymes like nicotinamide adenine dinucleotide phosphate (NADPH), NADPH oxidase (Nox), xanthine oxidase (XO), flavin oxidases from peroxisomes, and Cytochrome P450 present in the endoplasmic reticulum (ER) are the endogenous sources of ROS formation. Although, certain levels of ROS are necessary for maintaining metabolic pathways and cellular signaling that helps regulate various cellular activities like mitosis, proliferation, survival, migration, gene expression, and apoptosis (Liu et al., 2018). Further, transcription factors that regulate cellular responses to ROS stimulations are also activated in the presence of ROS. For instance, transcription factor NRF2-induced antioxidant mechanisms could be activated during high concentrations of ROS (Deshmukh et al., 2016)(Dai et al., 2018)(Xu et al., 2017a). Recently, Gao et al., 2018 in neuronal PC12 cells demonstrated that administration of trilobatin rescued from H_2O_2-induces cytotoxicity and prevents oxidative stress through regulating AMPK/NRF2/Sirt3 signaling cascade (Gao et al., 2018). Likewise, Sousa et al., 2020 demonstrated the involvement of NFκβ in ROS-mediated neurotoxic effects. In their study, the authors concluded that ocellatin peptides administration in hippocampal neurons reversed the lipopolysaccharides (LPS)-induced neurotoxicity through decreased ROS production and inhibition of NFκβ (Sousa et al., 2020).

The most common NDDs is AD, identified by gradual loss of neurons and accumulation of extracellular Aβ plaques and intracellular tau

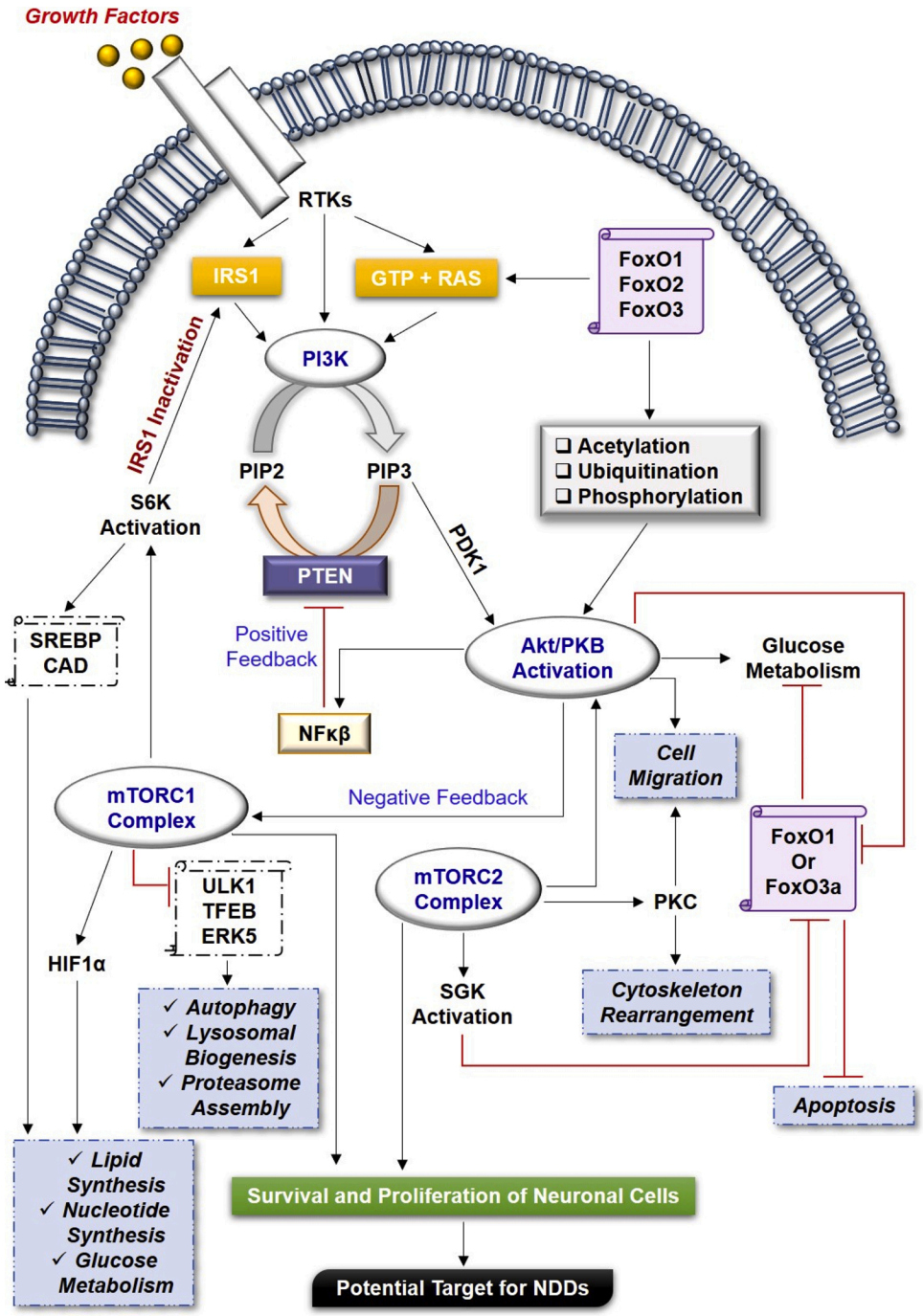


Fig. 4. AKT and mTORC mediated signaling in cell survival, proliferation, and growth. Growth factors activate the signaling cascade through RTKs, which in turn, through P13K forms PIP3 that activates AKT/PKB pathway. Further, this pathway supports its regulation by positive and negative feedback. In positive feedback, after AKT/PKB activation, NF-κB is produced, which inhibits PTEN, which can convert PIP3 to PIP2 as it can block AKT. In negative feedback, mTORC1 after AKT/PKB activation can induce S6K to inactivate IRS-1 so that no P13K is available and AKT/PKB remains in the inactivated form. AKT and mTORC2, through intercellular signaling and interaction with FOXO1 and FOXO3a, regulate glucose metabolism and apoptosis in the cell. mTORC2 through different signaling proteins maintain the survival and proliferation of cells. mTORC1 through its interaction with signaling proteins and AKT/PKB aid in cellular growth.

tangles (neurofibrillary tangles, NFT) (Weller and Budson, 2018). Different *in vivo* and *in vitro* studies demonstrated that imbalance redox signaling in neuronal cells is one of the majors contribute to neuronal cell death that leads to the development and progression of AD (Cheignon et al., 2018)(Simunkova et al., 2019)(Cioffi et al., 2019)(Cassidy et al., 2020)(González-Reyes et al., 2017). For instance, in mouse models of AD that express mutant amyloid precursor protein (APP) and presenilin-1 (PS-1), it is observed that excessive accumulation of H_2O_2 and lipid oxidation causes increased oxidative stress (Matsuoka et al., 2001). Similarly, the formation and aggregation of extracellular Aβ and enhanced phosphorylation of tau proteins due to oxidative stress and causes dangerous events of pathogenesis in AD (Zhao and Zhao, 2013). Likewise, in the AD mouse model with APP overexpression with deletion of cytoplasmic/zinc SOD in the Tg2576, which increases Aβ oligomerization and severe memory disruption (Murakami et al., 2011). There is evidence that cells with excessive tau proteins are prone to oxidative stress due to a decrease in the number of peroxisomes (Stamer et al., 2002). Similarly, the mouse model that expresses mutant (P301S and P301 L) tau protein has less activity of NADH-ubiquinone oxidoreductase and mitochondrial dysfunction, which are due to excessive ROS accumulation in neurons (Yoshiyama et al., 2007)(David et al., 2005)(Halverson, 2005). Further, mitochondrial dysfunction that is caused by Aβ blocks sufficient production of ATP and enhances ROS production in AD (Castellani et al., 2002). This is supported by observations that show reduced metabolism of energy in the brain during AD (Wang et al., 2014a). PD is another NDDs, which is due to a specific loss of neurons (dopaminergic neurons) in the *Substantia nigra pars compacta* (SNpc) and in the nigrostriatal DA (dopaminergic neuronal) pathway in

Table 3
Involvement of FOXO transcription factors in different neurodegenerative disorders.

Disease	Targeted Pathology	Clinical Significance of FOXO	References
Alzheimer's disease (AD)	Neurotoxic Aβ processing	Blocking of FoxO3 phosphorylation has a neuroprotective role	(Shi et al., 2016)
	Alternations in PI3K-Akt and MAPK/ERK signaling	Increased activity of FoxO causes alteration in autophagy and inflammation cascade	(Guillot et al., 2016)
	Regulate neuroinflammation and oxidative stress	Decreased FoxO1 activity increased AD pathology	(Yossef et al., 2020)
	Increased disease progression	Decreased FoxO1 activity enhanced the disease progression	(Wu et al., 2020)
	Reduces disease pathology	Overexpression of FoxO1 decreases Aβ production and tau phosphorylation	(Zhang et al., 2020)
	Regulates AchE inhibitors	High frequency of FoxO1 rs7981045 G/G genotype influence the clinical significance of AchE inhibitors in AD pathology	(Paroni et al., 2014)
	Impaired autophagy	FoxO6 promotes the expression of genes and proteins needed for autophagy	(Desai et al., 2020)
	Apoptosis modulator	Interaction of FoxO3a with BAF57 causes upregulation of PUMA, which leads to neuronal apoptosis	(Sanphui et al., 2020)
Parkinson's disease (PD)	Mitochondrial dysfunction and oxidative stress	Increased activity of FoxO induces activation of Wnt2, which protects the neuronal cell from oxidative stress	(Xia et al., 2020)
	Impaired autophagy	Activation of FoxO through Sirt1 activity causes enhanced autophagic degradation of α-synuclein	(Chen et al., 2019a)
	Oxidative damage	Inhibiting PI3K/Akt/FoxO signaling due to downregulation of SHC3 enhanced neuronal injury caused by oxidative stress	(Gong et al., 2018)
	Apoptosis regulation	FoxO regulates p53 activity through sirtuin expression, which induces apoptosis.	(Salimian et al., 2018)
	Stress response	Increased DAFa6/FoxO activates neuronal signaling pathways and rescues from oxidative stress	(Farina et al., 2017)
Huntington's disease (HD)	Disease progression	Genetic breakdown of dCBP promotes neuroprotection through FoxO dependent pathway	(Martin et al., 2021)
	Behavioral defects	Interaction between Poly Q proteins and FoxO proteins leads to dendritic pathology	(Kwon et al., 2018)
	Disease pathology	Increased FoxO3a levels contribute to disease progression	(Kannike et al., 2014)
Spinocerebellar ataxia (SA)	Decreased antioxidant capacity	through the positive feedback loop FoxO4 dependent SOD2 expression helps in reducing ROS levels	(de Assis et al., 2017)
	Stress response mechanism	Upregulation of FoxO4 causes increased activity of SOD2, which rescued from oxidative stress-induced neurotoxicity	(Araujo et al., 2011)
Frontotemporal dementia (FD)	Stress-induced protein misfolding	TDP-43 promotes activation of FoxO that helps in the expression of stress-resistant genes and protein homeostasis	(Zhang et al., 2014)
Amyotrophic lateral sclerosis (ALS)	Impaired integrity	FoxO1 activity causes increased expression of TJ protein that causes blood-spinal cord barrier breakdown	(Meister et al., 2015)
	TDP-43 proteinopathies	Mutation in the FoxO reversed the neuroprotection through α-Methyl-α-phenylsuccinimide	(Wong et al., 2018)
	Lewy body formation	FoxO3a, α-synuclein, and 14-3-3 protein may interact to form a complex that can promote neuronal survival	(Su et al., 2009)

the brain. (Keane et al., 2011). Apart from AD, oxidative stress play a significant role in the pathogenesis of PD (Guo et al., 2018)(Trist et al., 2019). For instance, studies have found that in the respiratory chain in *substantia nigra pars compacta* (SNpc) of patients with PD, and there is a decrease in the activity of Complex I due to which high levels of ROS are produced, and this leads to apoptosis (Jenner, 2003)(Schapira, 2008) (Franco-Iborra et al., 2016). Likewise, Singh et al., 2017 concluded that upregulation of Sirt1 rescued from α-synuclein aggregates-induced oxidative stress in SH-SY5Y cell culture model of PD by deactivating NFκβ and poly (ADP-ribose) polymerase (PARP) activity (Singh et al., 2017). Fig. 3 summarizes various ROS sources for the brain that can lead to high ROS levels in the brain, leading to neurodegeneration.

6. FOXO transcription factors: Implications in neuronal signaling and neurodegenerative disease

FOX transcription factors regulate neuronal and dendritic signaling by modulation of multiple signaling molecules involved in diverse physiological processes, namely stress resistance, metabolism, proliferation, immune response, and apoptosis (Santo and Paik, 2018). Further, the activation or deactivation of the FoxO family can be regulated by growth factors like IGF, which can support FoxO phosphorylation at the C-terminal side with the help of protein kinase B (Akt/PKB), due to which it cannot be retained in the nucleus and cannot perform its function as its nuclear localization signal (NLS) is masked (De Keizer et al., 2011). In addition, PTMs like phosphorylation through the JNK pathway can help the inactivation of FoxO in the presence of ROS (Rena

Table 4
List of several chemical and natural biomolecules that confer neuroprotection (*in vivo* and *in vitro*) through regulation of FoxO proteins.

Compound	Target	Interactors	Experimental Model	Mechanism	Reference
Vitis Vinifera	DAF-16/ FoxO	CAT, SODs, GSTs, and GPx	HT22 hippocampal neuronal cells	Increase antioxidant effect and upregulated stress resistance. Reduction in glutamate-induced neurotoxicity	(Duangjan et al., 2021)
S. macrocarpon L. (AE)	Jafrac1 and FOXO	AchE, Hsp70, Ace1, Dronc	*Drosophila melanogaster*	Decreases Alcl3-induced neurotoxicity by inhibiting oxidative stress through upregulation of FoxO and reduced activity of NF-Kκβ	(Ogunsuyi et al., 2021)
S. nigrum L. (BN)	Jafrac1 and FOXO	AchE, Hsp70, Ace1, Dronc	*Drosophila melanogaster*	Decreases $Alcl_3$-induced oxidative stress through upregulation of FoxO and decreased activity of NF-Kκβ	(Ogunsuyi et al., 2021)
Stigmasterol	FoxO3a and Sirtuin 1	Bcl-2	SH-SY5Y cells	Reversed H_2O_2-induced neurotoxicity through inhibiting oxidative stress. Administration of Stigmasterol upregulated FoxO3a activity and Bcl-2 protein expression in the neurons.	(Pratiwi et al., 2021)
Galangin	AP1 and FoxO1	MMP-9	Rat brain astrocytes cell culture model	Inhibits Pyk2/PDGFRβ/PI3K/Akt/mTOR/JNK1/JNK2 and p44/p42 MAPK cascade-dependent AP-1 and FoxO1 activities	(Yang et al., 2020)
10-HDA	FoxO1	NLRP3, TNFα, IL1β	C57BL/6 J mice and microglial Bv-2 cells	Activates FoxO1 nuclear translocation, which activates the mitophagy pathway and inhibits LPS-induced neurotoxicity. It also prevents neuroinflammation	(You et al., 2019)
Vigna unguiculata	FoxO	Sirtuin 1 and NOTCH	*Drosophila melanogaster*	Delay senescence and age-dependent neuronal degeneration	(Tripodi et al., 2020)
Diphenyl diselenide	DAF-16/ FoxO	Hsp16.2, SOD3, PhSe2	Caenorhabditis elegans model of HD	Activates antioxidant pathway through upregulation of FoxO proteins, which decreases protein aggregation and rescued neuronal cell death	(Bicca Obetine Baptista et al., 2020)
Urolithin A and actinonin	DAF-16/ FoxO	PINK1, PDR1, DCT1	Amyloid-β (Aβ) and tau *Caenorhabditis elegans* models of AD	Activates mitophagy through upregulation of DAF-16/FoxO	(Fang et al., 2019)
GLP-1RA	FoxO	Sirtuin 2	Diabetes-related dementia model	Rescued from metabolic memory-induced amyloid-β and tau pathologies through FoxO-dependent pathway	(Chen et al., 2019b)
α-Methyl-α-phenylsuccinimide	FoxO	TDP-43	*Caenorhabditis elegans*	Rescued protein aggregation through DAF-16/FoxO-dependent pathway	(Wong et al., 2018)
Rapamycin	FoxO	JNK, HSP	Drosophila melanogaster of spastic paraplegia	Prevents motor muscle degeneration through activation of JNK-FoxO stress response and decreases oxidative stress	(Xu et al., 2017b)
27-hydroxycholesterol	FoxO1	RTKN2, MAST4	APP-induced AD brain	Promotes neuroprotection through inhibition of MAST4-induced FoxO1/RTKN2 signaling pathway	(Gongol et al., 2017)
Ethosuximide	DAF-16/ FOXO	DNAJC5	*Caenorhabditis elegans*	Promotes neuroprotection through reduced cytotoxicity by activating DAF-16/FoxO pathway	(Chen et al., 2015)
Linagliptin	AKT1 and FOXO	___	*Caenorhabditis elegans*	Promotes antioxidant response through upregulation of FoxO and downregulation of Akt1	(Wongchai et al., 2015)
Liraglutide	AKT1 and FOXO	___	*Caenorhabditis elegans*	Promotes antioxidant response through upregulation of FoxO and downregulation of Akt1	(Wongchai et al., 2015)
GPI-1/GPI	DAF-16/ FOXO	IGF1, IIS	Parkinson's cell culture models	Prevents proteotoxicity and rescued neuronal cell death through upregulation of DAF-16/FoxO	(Knight et al., 2014)
Manganese	FoxO	SOD2, Akt	Parkinson's cell culture models	Manganese causes phosphorylation of FoxO, which protects astrocytes from oxidative stress through stress response mechanism and is independent of the insulin signaling cascade	(Exil et al., 2014)
TBHQ	FoxO3a	PI3K/Akt, NMDA	NMDA administered cortical neuronal culture	Induced Fox3a translocation, which inhibits its transcriptional activity and protects neurons from toxicity	(Bahia et al., 2012)
Organotellurium and organoselenium	DAF-16/ FoxO	DPTVP	*Caenorhabditis elegans*	Protects neuronal death through increased activity of DAF-16/FoxO and activates stress response mechanism	(Avila et al., 2012)
Estradiol	FoxO3a	PI3K/Akt	Hippocampal CA1 neurons of rodent	Rescued from ischemia-induced neuronal injury and cognitive dysfunction through prevention of GSK3β and FoxO phosphorylation	(Jover-Mengual et al., 2010)

et al., 2001). FoxO has a very conserved DNA binding domain and has a definite function in a particular tissue, and thus, the pathway that regulates FoxO in that cell will determine its role for the cell (Dansen, 2011). FoxO1a controls the expression of glucose-6-phosphatase and phosphoenolpyruvate carboxykinase, thereby regulating gluconeogenesis and glycogenolysis, and controls the checkpoints in the cell cycle by decreasing the cyclin D levels (De Keizer et al., 2010). Further, apoptosis is also affected by FoxO3a as it can regulate the expression of proapoptotic factors like Bim, p53 upregulated modulator of apoptosis (PUMA), and antiapoptotic factors like FLIP (De Keizer et al., 2011). Further, it was used to study the role of FoxO and related TFBEs in the regulation of oxidative stress. For instance, In the rat hippocampus, an increase in immunoreactivity of the dentate gyrus (DG) was observed, and it was visible from 15 to 60th day of exposure to ozone. In the western blot, it was evident that activation of FoxO3a takes place very significantly after 30 days and 60 days of ozone exposure that causes neurodegeneration (Bandmann et al., 1997). Fig. 4 shows AKT, mTORC1, and mTORC2 pathways that can regulate cell survival, proliferation, and growth.

Moreover, studies demonstrated the possible involvement of FoxO proteins in the pathogenesis and progression of NDDs, such as AD, PD, HD, ALS, dementia, and others. For instance, in AD pathogenesis, FoxO3 is considered the largest inducer of cell death in response to Aβ mediated oxidative stress. Administration of FoxO3 to the hippocampus and cortex region causes series of PTMs, which regulates FoxO3 activity and activates BIM that causes neuronal cell death. However, FoxO3-induced neuronal apoptosis could be reversed by the shRNA-induced knockdown of FoxO3a. Further, Aβ administration causes phosphorylation of MST1 at S209 followed by methylation and activation of protein arginine N-methyltransferase 1 (PRMT1) at Arg248 and Arg250, which promotes activation and nuclear translocation of FoxO3a that induces neuronal cell death (Maiese, 2016). Similarly, CDK5-mediated phosphorylation of FoxO3a in response to glutamate toxicity causes activation and nuclear translocation of FoxO3a that induces neuronal apoptosis through BIM activation. Genetic breakdown of FoxO3a or CDK5 prevents neuronal loss in response to glutamate toxicity. In addition, in the p25 overexpression mouse model of AD, enhanced activity of CDK5 and FoxO3a was observed that concludes the involvement of FoxO3a in neuronal cell death in response to Aβ toxicity (Shi et al., 2016). Recently, Chaubey

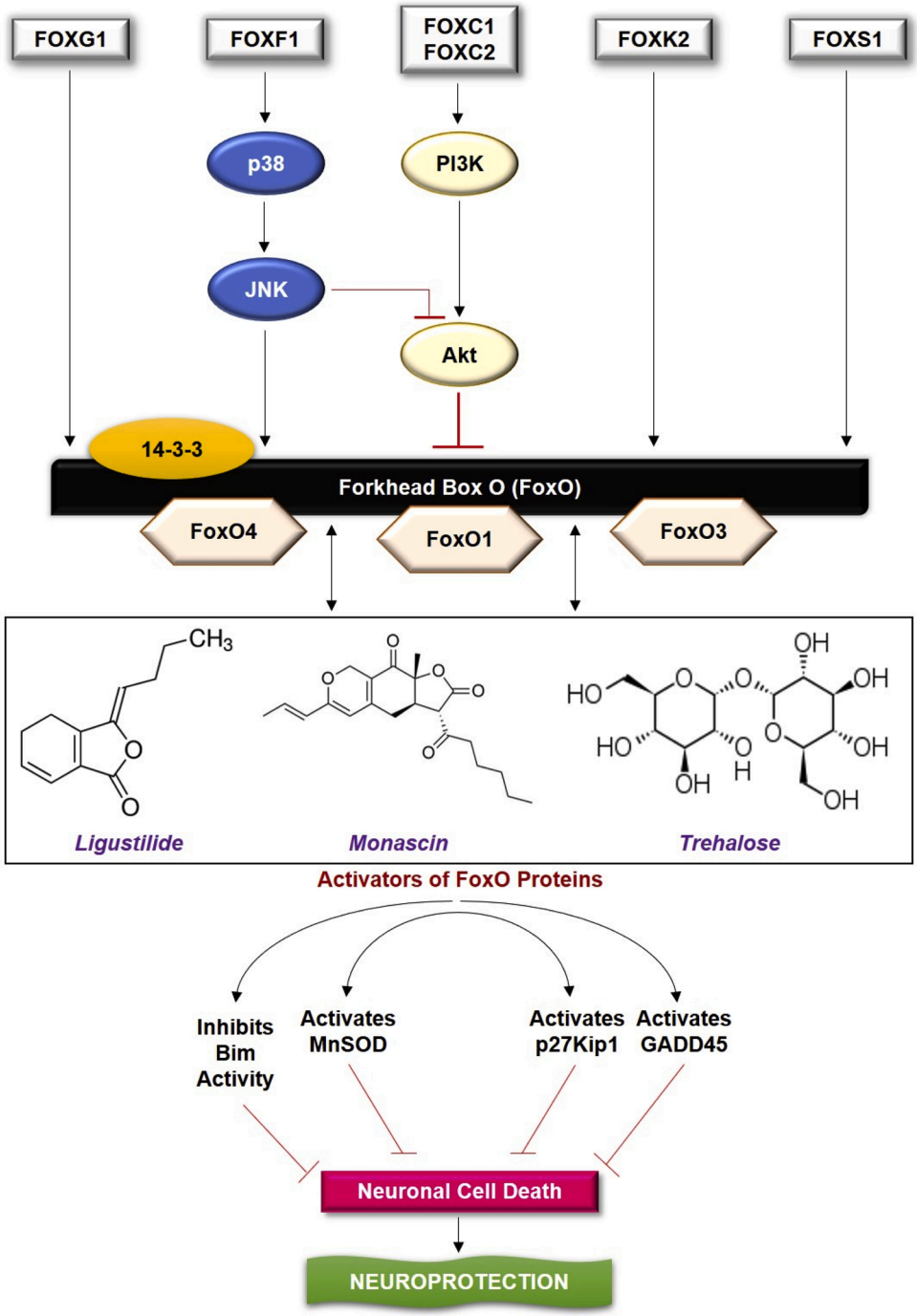


Fig. 5. FoxO precursors, namely FOXG1, FOXF1, FOXC1, FOXC2, FOXK2, and FOXS1, regulates the expression and activity of FoxO proteins, such as FoxO4, FoxO1, FoxO3. FOXG1, FOXK2, and FOXS1 directly activated the activity of FoxO proteins, whereas, FOXF1 interacts with p38, which activates p38/JNK complex and thus activates FoxO protein. Similarly, FOXC1 and FOXC2 activate PI3K followed by activation of Akt, which inhibits the activity of FoxO proteins. Several activators have been discovered, such as Ligustilide, Monascin, and Trehalose, which increase the expression of FoxO proteins and results in neuroprotection. Activation of FoxO protein inhibits the activity of Bim and later decreases neuronal cell apoptosis. In addition, increased activity of FoxO proteins causes activation of MnSOD, p27Kip1, and GADD45, which also inhibit neuronal cell death.

et al., 2020 demonstrated that FoxO was rescued from Aβ-induced oxidative stress and aging effects through regulating NRF2 upon allophycocyanin administration (Chaubey et al., 2020). In addition, the evidence for the involvement of FoxO transcription factor in the pathogenesis of AD came from the study that demonstrated that inhibition of miR-132/212 enhanced the expression of pro-apoptotic targets of FoxO3a, namely PTEN and p300, which caused cell death and contributed to the pathogenesis of AD (Wong et al., 2013). Similarly, accumulating evidence established the notion that FoxO transcription factors are also involved in the progression and pathogenesis of PD. For example, overexpression of FoxO3 causes neuronal apoptosis in dopaminergic neuronal cell culture. In addition, in response to oxidative stress and ER stress, the dominant-negative form of FoxO3 also induces neuronal cell apoptosis (Dumitriu et al., 2012). In conjugation with α-synuclein, FoxO causes acute toxicity that leads to neuronal apoptosis. The fact is established with the study that demonstrated that in α-synuclein overexpressing neurons, inhibition of FoxO3 promotes neuroprotection in response to oxidative stress suggesting that the toxic nature of FoxO3 is due to its transcriptional activity (Pino et al., 2014). Recently, Xia et al., 2020 concluded that overexpression of Wnt2 enhanced the expression of peroxisome proliferator-activated receptor-gamma coactivator 1-alpha (PGC-1α) and FoxO, which promotes neuroprotection in the PINK1B9 transgenic Drosophila PD model. Further, the authors concluded that the neuroprotective effect of Wnt2 overexpression is due to activation of the stress response mechanism through increased FoxO activity (Xia et al., 2020). In HD, loss of X-box-binding protein 1 (XBP1) causes FoxO1 accumulation in the nucleus, which decreases mHtt-induced neurotoxicity through enhanced

autophagic clearance of mHtt (Vidal et al., 2012). Similarly, enhanced expression of FoxO4 was observed in HD neurons, which modulate proteasome activity through Akt inhibition. The study concluded valuable therapeutic target of FoxO in HD pathology (Liu et al., 2017). However, further studies will require to extract the potential of FoxO in HD pathogenesis. Table 3 represents the implementation of FOXO transcription factors in different NDDs.

7. Pharmacological intervention targeting FoxO proteins in neurodegeneration

From the past several decades, attempts have been made in search of effective treatment against NDDs. However, due to complex disease pathogenesis and complex neuronal system, all therapeutic strategies go into vein. Recently, FoxO transcription factors have been demonstrated as an effective therapeutic target involved in neuronal signaling through insulin signaling and response to oxidative stress (Maiese, 2021). Several activators and inhibitors of FoxO transcription factors have been demonstrated *in vivo* and *in vitro* studies. For instance, melatonin regulates mitochondrial homeostasis and protects from Aβ-induced neurotoxicity through sirtuin1/FoxO1 pathway and PTEN/Akt/FoxO3a pathway, respectively (Zhao et al., 2018)(Jenwitheesuk et al., 2017). Similarly, derivatives of phenolic acids and carboxylic acids promote neuroprotection by increasing FoxO3a and sirtuin 1 expression by activating the stress response mechanism and inhibiting oxidative stress in neuroblastoma SH-SY5Y cultured cells (Gay et al., 2018). In addition, psammaplysene, dictyophora indusiata polysaccharide, and metformin also confer neuroprotection through upregulation of FoxO protein activity and activation of stress response mechanism *in vivo* and *in vitro* conditions (El-Ghaiesh et al., 2020)(Boccitto et al., 2017)(Zhang et al., 2016). Paiva et al., 2015 demonstrated that administration of Carqueja in a C. elegans model for AD protects from oxidative stress and Aβ-induced neurotoxicity through upregulation of SKN-1/Nrf and DAF-16/Foxo activity (Paiva et al., 2015). Further, intrinsic signaling plays an essential role in regulating FoxO activity in the treatment of NDDs. For example, L-NAME, a nitric oxide synthase inhibitor, inhibits the neuronal cell death in FoxO expressing aged flies by enhancing the phosphorylation of FoxO at S319 (Kanao et al., 2012). Further, Guitart et al., 2020 proposed a novel therapy, namely NeuroHeal, which promotes neuroprotection through autophagic degradation by increasing deacetylation and phosphorylation of FoxO3a in damaged neonatal motoneurons (Romeo-Guitart et al., 2020).

In addition, activation of FoxO transcription factors through enhanced p53 and LRRK2 activity confers neuroprotection in AD and PD experimental model, respectively (Akhter et al., 2014)(Kanao et al., 2010). Moreover, Bedada et al., 2021 concluded that exercise training rescued from cognitive impairment through upregulation of F-Box Protein 32 (FBXO32) and FoxO1 in the African American population (Bedada et al., 2021). Table 4 lists all the potential small molecules that confer neuroprotection through modulating FoxO activity. However, till date no clinical studies have been performed ton FoxO in the context of NDDs. Thus, further studies should be done to exploit the potential of FoxO transcription factors as a potential therapeutic target for the treatment of NDDs (Fig. 5).

8. Conclusions and future considerations

FOXO transcription factors have specific functions in neurons that are tightly regulated by signaling pathways and environmental stimulations. In this review, our focus was on the role of FOXO and related transcription factor binding events on NDDs. FoxO family proteins play an important role in regulating cell proliferation. Also, the expression of regulatory proteins of the cell cycle is controlled by FOXO3a, which in turn makes it control the cell cycle. In the future, this characteristic can be used to develop a regenerative medicinal approach. FoxOs have a significant regulatory function in learning, memory, and pro-longevity effects on neurons. The cognitive ability of the brain can be maintained with aging with the help of FoxOs, but this needs more evidence and research. The role of all isoforms of FoxO in different parts of the brain should be ascertained to propose a therapeutic method for NDDs. Furthermore, oxidative stress and excessive ROS presence in the neurons support the initiation and progression of NDDs like AD and PD. Due to ROS production in high levels, the neurons get damaged, and also it induces signals that result in apoptosis. The activated FoxOs can regulate oxidative stress and metabolic pathways to support neuronal survival. Also, it is necessary to maintain redox balance in neurons to treat or avoid NDDs. In conclusion, the FoxOs can be used as therapeutic targets to treat NDDs in the future. With the help of experimental procedures on mouse models, it is needed to investigate the use of FoxO in the regulation of oxidative stress, metabolic pathways, and maintenance of redox balance in neurons.

Author's contribution

PK Conceived this work. RG collected the data for revision. Fig. 5 and Table 4 along with data in other tables, restructuring the figures and points have been prepared by RG. VO and RG contributed equally to this work. All authors have role in the design, art work, analysis, results interpretation, and manuscript writing.

Funding

None.

Declaration of Competing Interest

The authors declare that there is no conflict of interests.

Acknowledgment

We would like to thanks the senior management of Delhi Technological University for their constant support and guidance.

References

Aghayeva, U., Bhattacharya, A., Sural, S., Jaeger, E., Churgin, M., Fang-Yen, C., Hobert, O., 2020. DAF-16/FoxO and DAF-12/VDR control cellular plasticity both cell-autonomously and via interorgan signaling. bioRxiv 6. https://doi.org/10.1101/2020.12.15.422925, 2020.12.15.422925.

Ahlenius, H., Chanda, S., Webb, A.E., Yousif, I., Karmazin, J., Prusiner, S.B., Brunet, A., Südhof, T.C., Wernig, M., 2016. FoxO3 regulates neuronal reprogramming of cells from postnatal and aging mice. Proc. Natl. Acad. Sci. 113, 8514–8519. https://doi.org/10.1073/pnas.1607079113.

Akhter, R., Sanphui, P., Biswas, S.C., 2014. The essential role of p53-up-regulated modulator of apoptosis (Puma) and its regulation by FoxO3a transcription factor in β-Amyloid-induced neuron death. J. Biol. Chem. 289, 10812. https://doi.org/10.1074/JBC.M113.519355.

Ali, Ammar, Shah, S.A., Zaman, N., Uddin, M.N., Khan, W., Ali, Abid, Riaz, M., Kamil, A., 2021. Vitamin D exerts neuroprotection via SIRT1/nrf-2/ NF-kB signaling pathways against D-galactose-induced memory impairment in adult mice. Neurochem. Int. 142, 104893 https://doi.org/10.1016/J.NEUINT.2020.104893.

Araujo, J., Breuer, P., Dieringer, S., Krauss, S., Dorn, S., Zimmermann, K., Pfeifer, A., Klockgether, T., Wuellner, U., Evert, B.O., 2011. FOXO4-dependent upregulation of superoxide dismutase-2 in response to oxidative stress is impaired in spinocerebellar ataxia type 3. Hum. Mol. Genet. 20, 2928–2941. https://doi.org/10.1093/hmg/ddr197.

Audesse, A.J., Dhakal, S., Hassell, L.-A., Gardell, Z., Nemtsova, Y., Webb, A.E., 2019. FOXO3 directly regulates an autophagy network to functionally regulate proteostasis in adult neural stem cells. PLoS Genet. 15, e1008097 https://doi.org/10.1371/JOURNAL.PGEN.1008097.

Avila, D.S., Benedetto, A., Au, C., Manarin, F., Erikson, K., Soares, F.A., Rocha, J.B.T., Aschner, M., 2012. Organotellurium and Organoselenium Compounds attenuate Mn-induced toxicity in C. elegans by preventing oxidative stress. Free Radic. Biol. Med. 52, 1903. https://doi.org/10.1016/J.FREERADBIOMED.2012.02.044.

Bahia, P.K., Pugh, V., Hoyland, K., Hensley, V., Rattray, M., Williams, R.J., 2012. Neuroprotective effects of phenolic antioxidant tBHQ associate with inhibition of FoxO3a nuclear translocation and activity. J. Neurochem. 123, 182. https://doi.org/10.1111/J.1471-4159.2012.07877.X.

Bandmann, O., Sweeney, M.G., Daniel, S.E., Marsden, C.D., Wood, N.W., 1997. Mitochondrial DNA polymorphisms in pathologically proven Parkinson's disease. J. Neurol. 244, 262–265. https://doi.org/10.1007/s004150050082.

Bedada, F.B., Ntekim, O.E., Nwulia, E.O., Fungwe, T.V., Nadarajah, S.R., Obisesan, T.O., 2021. Exercise training-increased FBXO32 and FOXO1 in a gender-dependent manner in mild cognitively impaired African Americans: GEMS-1 study. Front. Aging Neurosci. 13 https://doi.org/10.3389/FNAGI.2021.641758.

Bernard, M.Law, Guest, Amy L., Pullen, Matthew W.J., Michael, S., Perkinton, Robert J. Williams, 2018. Increased Foxo3a nuclear translocation and activity is an early neuronal response to βγ-secretase-mediated processing of the amyloid-β protein precursor: utility of an AβPP-GAL4 reporter assay. J. Alzheimers Dis. 61, 673–688. https://doi.org/10.3233/JAD-170393.

Bicca Obetine Baptista, F., Arantes, L.P., Machado, M.L., da Silva, A.F., Marafiga Cordeiro, L., da Silveira, T.L., Soares, F.A.A., 2020. Diphenyl diselenide protects a Caenorhabditis elegans model for Huntington's disease by activation of the antioxidant pathway and a decrease in protein aggregation. Metallomics 12, 1142–1158. https://doi.org/10.1039/D0MT00074D.

Biggs III, W.H., Cavenee Karen C, W.K., 2001. Identification and characterization of members of the FKHR (FOX O) subclass of winged-helix transcription factors in the mouse. Mamm. Genome 12, 416–425. https://doi.org/10.1007/s003350020002.

Blesa, J., Przedborski, S., 2014. Parkinsonâ€TMs disease: animal models and dopaminergic cell vulnerability. Front. Neuroanat. 8 https://doi.org/10.3389/fnana.2014.00155.

Boccitto, M., Lee, N., Sakamoto, S., Spruce, L.A., Handa, H., Clardy, J., Seeholzer, S.H., Kalb, R.G., 2017. The neuroprotective marine compound psammaplysene a binds the RNA-Binding protein HNRNPK. Mar. Drugs 15. https://doi.org/10.3390/MD15080246. Page 246 15, 246. 2017.

Brenkman, A.B., Keizer, P.L.Jde, Broek, N.J.Fvanden, Jochemsen, A.G., Boudewijn, M., Burgering, T., 2008. Mdm2 induces mono-ubiquitination of FOXO4. PLoS One 3, e2819. https://doi.org/10.1371/JOURNAL.PONE.0002819.

Brown, A.K., Webb, A.E., 2018. Regulation of FOXO factors in mammalian cells. Curr. Top. Dev. Biol. 127, 165–192. https://doi.org/10.1016/BS.CTDB.2017.10.006.

Cai, H., Wang, Y., McCarthy, D., Wen, H., Borchelt, D.R., Price, D.L., Wong, P.C., 2001. BACE1 is the major β-secretase for generation of Aβ peptides by neurons. Nat. Neurosci. 4, 233–234. https://doi.org/10.1038/85064.

Cassidy, L., Fernandez, F., Johnson, J.B., Naiker, M., Owoola, A.G., Broszczak, D.A., 2020. Oxidative stress in alzheimer's disease: a review on emergent natural polyphenolic therapeutics. Complement. Ther. Med. 49, 102294 https://doi.org/10.1016/J.CTIM.2019.102294.

Castellani, R., Hirai, K., Aliev, G., Drew, K.L., Nunomura, A., Takeda, A., Cash, A.D., Obrenovich, M.E., Perry, G., Smith, M.A., 2002. Role of mitochondrial dysfunction in Alzheimer's disease. J. Neurosci. Res. 70, 357–360. https://doi.org/10.1002/jnr.10389.

Chang, G., Chen, Y., Zhang, H., Zhou, W., 2019. Trans sodium crocetinate alleviates ischemia/reperfusion-induced myocardial oxidative stress and apoptosis via the SIRT3/FOXO3a/SOD2 signaling pathway. Int. Immunopharmacol. 71, 361–371. https://doi.org/10.1016/J.INTIMP.2019.03.056.

Chaubey, M.G., Patel, S.N., Rastogi, R.P., Madamwar, D., Singh, N.K., 2020. Cyanobacterial pigment protein allophycocyanin exhibits longevity and reduces Aβ-mediated paralysis in C. elegans : complicity of FOXO and NRF2 ortholog DAF-16 and SKN-1. 3 Biotech 10, 1–11. https://doi.org/10.1007/S13205-020-02314-1, 2020 108.

Cheignon, C., Tomas, M., Bonnefont-Rousselot, D., Faller, P., Hureau, C., Collin, F., 2018. Oxidative stress and the amyloid beta peptide in Alzheimer's disease. Redox Biol. 14, 450–464. https://doi.org/10.1016/J.REDOX.2017.10.014.

Chen, X., McCueF, H.V., Wong, S.Q., Kashyap, S.S., Kraemer, B.C., Barclay, J.W., Burgoyne, R.D., Morgan, A., 2015. Ethosuximide ameliorates neurodegenerative disease phenotypes by modulating DAF-16/FOXO target gene expression. Mol. Neurodegener. 10 https://doi.org/10.1186/S13024-015-0046-3.

Chen, C., Xia, B., Tang, L., Wu, W., Tang, J., Liang, Y., Yang, H., Zhang, Z., Lu, Y., Chen, G., Yang, Y., Zhao, Y., 2019a. Echinacoside protects against MPTP/MPP+-induced neurotoxicity via regulating autophagy pathway mediated by Sirt1. Metab. Brain Dis. 34, 203. https://doi.org/10.1007/S11011-018-0330-3.

Chen, S., Tang, Q., Wang, Y., Xu, Z., Chen, S.T., Sun, Y., Yao, W.B., Gao, X.D., 2019b. Evidence of metabolic memory-induced neurodegeneration and the therapeutic effects of glucagon-like peptide-1 receptor agonists via Forkhead box class O. Biochim. Biophys. Acta - Mol. Basis Dis. 1865, 371–377. https://doi.org/10.1016/J.BBADIS.2018.11.012.

Cioffi, F., Adam, R.H.I., Broersen, K., 2019. Molecular mechanisms and genetics of oxidative stress in Alzheimer's disease. J. Alzheimers Dis. 72, 981–1017. https://doi.org/10.3233/JAD-190863.

Dai, Y., Zhang, H., Zhang, J., Yan, M., 2018. Isoquercetin attenuates oxidative stress and neuronal apoptosis after ischemia/reperfusion injury via Nrf2-mediated inhibition of the NOX4/ROS/NF-κB pathway. Chem. Biol. Interact. 284, 32–40. https://doi.org/10.1016/J.CBI.2018.02.017.

Daitoku, H., Sakamaki, Jichi, Fukamizu, A., 2011. Regulation of FoxO transcription factors by acetylation and protein–protein interactions. Biochim. Biophys. Acta - Mol. Cell Res. 1813, 1954–1960. https://doi.org/10.1016/J.BBAMCR.2011.03.001.

Dansen, T.B., 2011. Forkhead Box O transcription factors: key players in redox signaling. Antioxidants Redox Signal. https://doi.org/10.1089/ars.2010.3778.

David, D.C., Hauptmann, S., Scherping, I., Schuessel, K., Keil, U., Rizzu, P., Ravid, R., Dröse, S., Brandt, U., Müller, W.E., Eckert, A., Götz, J., 2005. Proteomic and functional analyses reveal a mitochondrial dysfunction in P301L tau transgenic mice. J. Biol. Chem. 280, 23802–23814. https://doi.org/10.1074/jbc.M500356200.

Dawson, T.M., Ko, H.S., Dawson, V.L., 2010. Genetic animal models of Parkinson's disease. Neuron 66, 646–661. https://doi.org/10.1016/j.neuron.2010.04.034.

de Assis, A.M., Saute, J.A.M., Longoni, A., Haas, C.B., Torrez, V.R., Brochier, A.W., Souza, G.N., Furtado, G.V., Gheno, T.C., Russo, A., Monte, T.L., Castilhos, R.M., Schumacher-Schuh, A., D'Avila, R., Donis, K.C., de Mello Rieder, C.R., Souza, D.O., Camey, S., Leotti, V.B., Jardim, L.B., Portela, L.V., 2017. Peripheral oxidative stress biomarkers in spinocerebellar Ataxia type 3/Machado–Joseph disease. Front. Neurol. 8 https://doi.org/10.3389/fneur.2017.00485.

De Keizer, P.L.J., Packer, L.M., Szypowska, A.A., Riedl-Polderman, P.E., Van Den Broek, N.J.F., De Bruin, A., Dansen, T.B., Marais, R., Brenkman, A.B., Burgering, B. M.T., 2010. Activation of forkhead box O transcription factors by oncogenic BRAF promotes p21cip1-dependent senescence. Cancer Res. 70, 8526–8536. https://doi.org/10.1158/0008-5472.CAN-10-1563.

De Keizer, P.L.J., Burgering, B.M.T., Dansen, T.B., 2011. Forkhead box O as a sensor, mediator, and regulator of redox signaling. Antioxidants Redox Signal. https://doi.org/10.1089/ars.2010.3403.

Deng, A., Ma, L., Zhou, X., Wang, X., Wang, S., Chen, X., 2020. FoxO3 Transcription Factor Promotes Autophagy after Oxidative Stress Injury in HT22 Cells, 99, pp. 627–634. https://doi.org/10.1139/cjpp-2020-0448 https://doi.org/10.1139/CJPP-2020-0448.

Desai, S., Pansare, P., Sainani, S., Doke, R., Bhalchim, V., Rode, K., 2020. Foxo6 – a novel target for Parkinson's disease. Biomed. Pharmacol. J. 13, 367–381. https://doi.org/10.13005/bpj/1897.

Deshmukh, P., Unni, S., Krishnappa, G., Padmanabhan, B., 2016. The Keap1–Nrf2 pathway: promising therapeutic target to counteract ROS-mediated damage in cancers and neurodegenerative diseases. Biophys. Rev. 2016, 41–56. https://doi.org/10.1007/S12551-016-0244-4, 91 9.

Duangjan, C., Rangsinth, P., Zhang, S., Gu, X., Wink, M., Tencomnao, T., 2021. Vitis vinifera leaf extract protects against glutamate-induced oxidative toxicity in HT22 hippocampal neuronal cells and increases stress resistance properties in Caenorhabditis elegans. Front. Nutr. 8, 634100 https://doi.org/10.3389/FNUT.2021.634100.

Dumitriu, A., Latourelle, J.C., Hadzi, T.C., Pankratz, N., Garza, D., Miller, J.P., Vance, J. M., Foroud, T., Beach, T.G., Myers, R.H., 2012. Gene expression profiles in parkinson disease prefrontal cortex implicate FOXO1 and genes under its transcriptional regulation. PLoS Genet. 8, 1002794 https://doi.org/10.1371/JOURNAL.PGEN.1002794.

El-Ghaiesh, S.H., Bahr, H.I., Ibrahiem, A.T., Ghorab, D., Alomar, S.Y., Farag, N.E., Zaitone, S.A., 2020. Metformin protects from rotenone–Induced nigrostriatal neuronal death in adult mice by activating AMPK-FOXO3 signaling and mitigation of angiogenesis. Front. Mol. Neurosci. 0, 84. https://doi.org/10.3389/FNMOL.2020.00084.

Exil, V., Ping, L., Yu, Y., Chakraborty, S., Caito, S.W., Wells, K.S., Karki, P., Lee, E., Aschner, M., 2014. Activation of MAPK and FoxO by manganese (Mn) in rat neonatal primary astrocyte cultures. PLoS One 9, e94753. https://doi.org/10.1371/JOURNAL.PONE.0094753.

Fang, E.F., Hou, Y., Palikaras, K., Adriaanse, B.A., Kerr, J.S., Yang, B., Lautrup, S., Hasan-Olive, M.M., Caponio, D., Dan, X., Rocktäschel, P., Croteau, D.L., Akbari, M., Greig, N.H., Fladby, T., Nilsen, H., Cader, M.Z., Mattson, M.P., Tavernarakis, N., Bohr, V.A., 2019. Mitophagy inhibits amyloid-β and tau pathology and reverses cognitive deficits in models of Alzheimer's disease. Nat. Neurosci. 223 22, 401–412. https://doi.org/10.1038/s41593-018-0332-9, 2019.

Farina, F., Lambert, E., Commeau, L., Lejeune, F.-X., Roudier, N., Fonte, C., Parker, J.A., Boddaert, J., Verny, M., Baulieu, E.-E., Neri, C., 2017. The stress response factor daf-16 /FOXO is required for multiple compound families to prolong the function of neurons with Huntington's disease. Sci. Reports 1–15. https://doi.org/10.1038/s41598-017-04256-w, 2017 71 7.

Ferber, E.C., Peck, B., Delpuech, O., Bell, G.P., East, P., Schulze, A., 2012. FOXO3a regulates reactive oxygen metabolism by inhibiting mitochondrial gene expression. Cell Death Differ. 19, 968–979. https://doi.org/10.1038/cdd.2011.179.

Ferguson, D., Shao, N., Heller, E., Feng, J., Neve, R., Kim, H.-D., Call, T., Magazu, S., Shen, L., Nestler, E.J., 2015. SIRT1-FOXO3a regulate cocaine actions in the nucleus accumbens. J. Neurosci. 35, 3100–3111. https://doi.org/10.1523/JNEUROSCI.4012-14.2015.

Fleming, S.M., Fernagut, P.-O., Chesselet, M.-F., 2005. Genetic mouse models of parkinsonism: strengths and limitations. NeuroRX 2, 495–503. https://doi.org/10.1602/neurorx.2.3.495.

Franco-Iborra, S., Vila, M., Perier, C., 2016. The parkinson disease mitochondrial hypothesis. Neurosci. 22, 266–277. https://doi.org/10.1177/1073858415574600.

Gandhi, S., Abramov, A.Y., 2012. Mechanism of oxidative stress in neurodegeneration. Oxid. Med. Cell. Longev. 2012, 1–11. https://doi.org/10.1155/2012/428010.

Gao, J., Liu, S., Xu, F., Liu, Y., Lv, C., Deng, Y., Shi, J., Gong, Q., 2018. Trilobatin protects against oxidative injury in neuronal PC12 cells through regulating mitochondrial ROS homeostasis mediated by AMPK/Nrf2/Sirt3 signaling pathway. Front. Mol. Neurosci. 11, 267. https://doi.org/10.3389/FNMOL.2018.00267.

Gay, N.H., Phopin, K., Suwanjang, W., Songtawee, N., Ruankham, W., Wongchitrat, P., Prachayasittikul, S., Prachayasittikul, V., 2018. Neuroprotective effects of phenolic and carboxylic acids on oxidative stress-induced toxicity in human neuroblastoma SH-SY5Y cells. Neurochem. Res. 2018, 619–636. https://doi.org/10.1007/S11064-017-2463-X, 433 43.

Gitler, A.D., Dhillon, P., Shorter, J., 2017. Neurodegenerative disease: models, mechanisms, and a new hope. DMM Dis. Model. Mech. https://doi.org/10.1242/dmm.030205.

Goldbraikh, D., Neufeld, D., Eid-Mutlak, Y., Lasry, I., Gilda, J.E., Parnis, A., Cohen, S., 2020. USP1 deubiquitinates Akt to inhibit PI3K-Akt-FoxO signaling in muscle during prolonged starvation. EMBO Rep. 21, e48791 https://doi.org/10.15252/EMBR.201948791.

Gong, J., Zhang, L., Zhang, Q., Li, X., Xia, X.-J., Liu, Y.-Y., Yang, Q.-S., 2018. Lentiviral vector-mediated SHC3 silencing exacerbates oxidative stress injury in nigral dopamine neurons by regulating the PI3K-AKT-FoxO signaling pathway in rats with Parkinson's disease. Cell. Physiol. Biochem. 49, 971–984. https://doi.org/10.1159/000493228.

Gongol, B., Marin, T.L., Jeppson, J.D., Mayagoitia, K., Shin, S., Sanchez, N., Kirsch, W.M., Vinters, H.V., Wilson, C.G., Ghribi, O., Soriano, S., 2017. Cellular hormetic response to 27-hydroxycholesterol promotes neuroprotection through AICD induction of MAST4 abundance and kinase activity. Sci. Reports 2017, 1–11. https://doi.org/10.1038/s41598-017-13933-9, 71 7.

González-Reyes, R.E., Nava-Mesa, M.O., Vargas-Sánchez, K., Ariza-Salamanca, D., Mora-Muñoz, L., 2017. Involvement of astrocytes in alzheimer's disease from a neuroinflammatory and oxidative stress perspective. Front. Mol. Neurosci. 10, 427. https://doi.org/10.3389/FNMOL.2017.00427.

Gray, M., Shirasaki, D.I., Cepeda, C., Andre, V.M., Wilburn, B., Lu, X.-H., Tao, J., Yamazaki, I., Li, S.-H., Sun, Y.E., Li, X.-J., Levine, M.S., Yang, X.W., 2008. Full-length human mutant huntingtin with a stable polyglutamine repeat can elicit progressive and selective neuropathogenesis in BACHD mice. J. Neurosci. 28, 6182–6195. https://doi.org/10.1523/JNEUROSCI.0857-08.2008.

Guillot, F., Kemppainen, S., Lavasseur, G., Miettinen, P.O., Laroche, S., Tanila, H., Davis, S., 2016. Brain-specific basal and novelty-induced alternations in PI3K-Akt and MAPK/ERK signaling in a middle-aged AβPP/PS1 mouse model of Alzheimer's disease. J. Alzheimers Dis. 51, 1157–1173. https://doi.org/10.3233/JAD-150926.

Guo, J., Zhao, X., Li, Y., Li, G., Liu, X., 2018. Damage to dopaminergic neurons by oxidative stress in Parkinson's disease (Review). Int. J. Mol. Med. 41, 1817–1825. https://doi.org/10.3892/IJMM.2018.3406.

Halverson, R.A., 2005. Tau protein is cross-linked by transglutaminase in P301L tau transgenic mice. J. Neurosci. 25, 1226–1233. https://doi.org/10.1523/JNEUROSCI.3263-04.2005.

Harada, A., Oguchi, K., Okabe, S., Kuno, J., Terada, S., Ohshima, T., Sato-Yoshitake, R., Takei, Y., Noda, T., Hirokawa, N., 1994. Altered microtubule organization in small-calibre axons of mice lacking tau protein. Nature 369, 488–491. https://doi.org/10.1038/369488a0.

Heemels, M.T., 2016. Neurodegenerative diseases. Nature. https://doi.org/10.1038/539179a.

Hesp, K., Smant, G., Kammenga, J.E., 2015. Caenorhabditis elegans DAF-16/FOXO transcription factor and its mammalian homologs associate with age-related disease. Exp. Gerontol. 72, 1–7. https://doi.org/10.1016/J.EXGER.2015.09.006.

Hoekman, M.F.M., Jacobs, F.M.J., Smidt, M.P., Burbach, J.P.H., 2006. Spatial and temporal expression of FoxO transcription factors in the developing and adult murine brain. Gene Expr. Patterns 6, 134–140. https://doi.org/10.1016/j.modgep.2005.07.003.

Hong, Y.K., Lee, S., Park, S.H., Lee, J.H., Han, S.Y., Kim, S.T., Kim, Y.K., Jeon, S., Koo, B. S., Cho, K.S., 2012. Inhibition of JNK/dFOXO pathway and caspases rescues neurological impairments in Drosophila alzheimer's disease model. Biochem. Biophys. Res. Commun. 419, 49–53. https://doi.org/10.1016/J.BBRC.2012.01.122.

Horst, Avander, Tertoolen, L.G.J., Vries-Smits, L.M.Mde, Frye, R.A., Medema, R.H., Burgering, B.M.T., 2004. FOXO4 is acetylated upon peroxide stress and deacetylated by the longevity protein hSir2SIRT1*. J. Biol. Chem. 279, 28873–28879. https://doi.org/10.1074/JBC.M401138200.

Huang, H., Tindall, D.J., 2011. Regulation of FOXO protein stability via ubiquitination and proteasome degradation. Biochim. Biophys. Acta - Mol. Cell Res. 1813, 1961–1964. https://doi.org/10.1016/J.BBAMCR.2011.01.007.

Hwang, S., Song, S., Hong, Y.K., Choi, G., Suh, Y.S., Han, S.Y., Lee, M., Park, S.H., Lee, J. H., Lee, S., Bang, S.M., Jeong, Y., Chung, W.-J., Lee, I.-S., Jeong, G., Chung, J., Cho, K.S., 2013. Drosophila DJ-1 decreases neural sensitivity to stress by negatively regulating daxx-like protein through dFOXO. PLoS Genet. 9, e1003412 https://doi.org/10.1371/journal.pgen.1003412.

Hwang, I., Oh, H., Santo, E., Kim, D.-Y., Chen, J.W., Bronson, R.T., Locasale, J.W., Na, Y., Lee, J., Reed, S., Toth, M., Yu, W.H., Muller, F.L., Paik, J., 2018. FOXO protects against age-progressive axonal degeneration. Aging Cell 17, e12701. https://doi.org/10.1111/ACEL.12701.

Jacobs, F.M.J., van der Heide, L.P., Wijchers, P.J.E.C., Burbach, J.P.H., Hoekman, M.F. M., Smidt, M.P., 2003. FoxO6, a novel member of the FoxO class of transcription factors with distinct shuttling dynamics. J. Biol. Chem. 278, 35959–35967. https://doi.org/10.1074/jbc.M302804200.

Jenner, P., 2003. Oxidative stress in Parkinson's disease. Ann. Neurol. 53, S26–S38. https://doi.org/10.1002/ana.10483.

Jenwitheesuk, A., Boontem, P., Wongchitrat, P., Tocharus, J., Mukda, S., Govitrapong, P., 2017. Melatonin regulates the aging mouse hippocampal homeostasis via the sirtuin1-FOXO1 pathway. EXCLI J. 16, 340. https://doi.org/10.17179/EXCLI2016-852.

Jiang, C., Wan, X., He, Y., Pan, T., Jankovic, J., Le, W., 2005. Age-dependent dopaminergic dysfunction in Nurr1 knockout mice. Exp. Neurol. 191, 154–162. https://doi.org/10.1016/j.expneurol.2004.08.035.

Jover-Mengual, T., Miyawaki, T., Latuszek, A., Alborch, E., Zukin, R.S., Etgen, A.M., 2010. Acute estradiol protects CA1 neurons from ischemia-induced apoptotic cell death via the PI3K/Akt pathway. Brain Res. 1321, 1–12. https://doi.org/10.1016/J.BRAINRES.2010.01.046.

Kaestner, K.H., Knöchel, W., Martínez, D.E., 2000. Unified nomenclature for the winged helix/forkhead transcription factors. Genes Dev. https://doi.org/10.1101/gad.14.2.142.

Kanao, T., Venderova, K., Park, D.S., Unterman, T., Lu, B., Imai, Y., 2010. Activation of FoxO by LRRK2 induces expression of proapoptotic proteins and alters survival of postmitotic dopaminergic neuron in Drosophila. Hum. Mol. Genet. 19, 3747–3758. https://doi.org/10.1093/HMG/DDQ289.

Kanao, T., Sawada, T., Davies, S.-A., Ichinose, H., Hasegawa, K., Takahashi, R., Hattori, N., Imai, Y., 2012. The nitric oxide-cyclic GMP pathway regulates FoxO and alters dopaminergic neuron survival in Drosophila. PLoS One 7, 30958. https://doi.org/10.1371/JOURNAL.PONE.0030958.

Kannike, K., Sepp, M., Zuccato, C., Cattaneo, E., Timmusk, T., 2014. Forkhead transcription factor FOXO3a levels are increased in huntington disease because of overactivated positive autofeedback loop *. J. Biol. Chem. 289, 32845–32857. https://doi.org/10.1074/JBC.M114.612424.

Katila, N., Bhurtel, S., Park, P.H., Hong, J.T., Choi, D.Y., 2020. Activation of AMPK/aPKCζ/CREB pathway by metformin is associated with upregulation of GDNF and dopamine. Biochem. Pharmacol. 180, 114193 https://doi.org/10.1016/J.BCP.2020.114193.

Kauffman, A.L., Ashraf, J.M., Corces-Zimmerman, M.R., Landis, J.N., Murphy, C.T., 2010. Insulin signaling and dietary restriction differentially influence the decline of learning and memory with age. PLoS Biol. 8, e1000372 https://doi.org/10.1371/journal.pbio.1000372.

Keane, P.C., Kurzawa, M., Blain, P.G., Morris, C.M., 2011. Mitochondrial dysfunction in parkinson's disease. Parkinsons Dis. 2011, 1–18. https://doi.org/10.4061/2011/716871.

Kim, S.Y., Webb, A.E., 2017. Neuronal functions of FOXO/DAF-16. Nutr. Heal. Aging. https://doi.org/10.3233/NHA-160009.

Kim, D.-Y., Hwang, I., Muller, F.L., Paik, J.-H., 2015a. Functional regulation of FoxO1 in neural stem cell differentiation. Cell Death Differ. 22, 2034. https://doi.org/10.1038/CDD.2015.123.

Kim, G.H., Kim, J.E., Rhie, S.J., Yoon, S., 2015b. The role of oxidative stress in neurodegenerative diseases. Exp. Neurobiol. https://doi.org/10.5607/en.2015.24.4.325.

Knight, A.L., Yan, X., Hamamichi, S., Ajjuri, R.R., Mazzulli, J.R., Zhang, M.W., Daigle, J. G., Zhang, S., Borom, A.R., Roberts, L.R., Lee, S.K., DeLeon, S.M., Viollet-Djelassi, C., Krainc, D., O'Donnell, J.M., Caldwell, K.A., Caldwell, G.A., 2014. The glycolytic enzyme, GPI-1, is a functionally conserved modifier of DOPAMINERIGIC neurodegeneration in Parkinson's models[GC1]. Cell Metab. 20, 145. https://doi.org/10.1016/J.CMET.2014.04.017.

Kops, G.J.P.L., Ruiter, N.Dde, De Vries-Smits, A.M.M., Powell, D.R., Bos, J.L., Burgering, B.M.T., 1999. Direct control of the Forkhead transcription factor AFX by protein kinase B. Nature 398, 630–634. https://doi.org/10.1038/19328.

Kwon, M.J., Han, M.H., Bagley, J.A., Hyeon, D.Y., Ko, B.S., Lee, Y.M., Cha, I.J., Kim, S.Y., Kim, D.Y., Kim, H.M., Hwang, D., Lee, S.B., Jan, Y.N., 2018. Coiled-coil structure-dependent interactions between polyQ proteins and Foxo lead to dendrite pathology and behavioral defects. Proc. Natl. Acad. Sci. 115, E10748–E10757. https://doi.org/10.1073/PNAS.1807206115.

Lau, C., Koty, Z., Nalbantoglu, J., 2009. Differential response of glioma cells to FOXO1-directed therapy. Cancer Res. 69.

Law, B.M., Guest, A.L., Pullen, M.W.J., Perkinton, M.S., Williams, R.J., 2018. Increased Foxo3a nuclear translocation and activity is an early neuronal response to βγ-secretase-Mediated processing of the Amyloid-β protein precursor: utility of an AβPP-GAL4 reporter assay. J. Alzheimers Dis. 61, 673–688. https://doi.org/10.3233/JAD-170393.

Libina, N., Berman, J.R., Kenyon, C., 2003. Tissue-specific activities of C. Elegans DAF-16 in the regulation of lifespan. Cell 115, 489–502. https://doi.org/10.1016/S0092-8674(03)00889-4.

Lin, M.T., Beal, M.F., 2006. Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. Nature 443, 787–795. https://doi.org/10.1038/nature05292.

Lin, S.-P., Hu, J., Wei, J.-X., Ye, S., Bu, J., Xu, W., Chen, S., Wu, Y., Wu, G., Zhu, L., Lin, P.-Y., Chen, X.-H., 2020. Silencing of circFoxO3 protects HT22 cells from glutamate-induced oxidative injury via regulating the mitochondrial apoptosis pathway. Cell. Mol. Neurobiol. 2020, 1231–1242. https://doi.org/10.1007/S10571-020-00817-2, 407 40.

Liu, Y., Qiao, F., Leiferman, P.C., Ross, A., Schlenker, E.H., Wang, H., 2017. FOXOs modulate proteasome activity in human-induced pluripotent stem cells of Huntington's disease and their derived neural cells. Hum. Mol. Genet. 26, 4416. https://doi.org/10.1093/HMG/DDX327.

Liu, Z., Ren, Z., Zhang, J., Chuang, C.-C., Kandaswamy, E., Zhou, T., Zuo, L., 2018. Role of ROS and nutritional antioxidants in human diseases. Front. Physiol. 9, 477. https://doi.org/10.3389/FPHYS.2018.00477.

Lu, T., Pan, Y., Kao, S.-Y., Li, C., Kohane, I., Chan, J., Yankner, B.A., 2004. Gene regulation and DNA damage in the ageing human brain. Nature 429, 883–891. https://doi.org/10.1038/nature02661.

Ma, Z., Zhang, W., Fan, W., Wu, Y., Zhang, M., Xu, J., Li, W., Sun, L., Liu, Wenjun, Liu, Wei, 2021. Forkhead box O1-mediated ubiquitination suppresses RIG-I-mediated antiviral immune responses. Int. Immunopharmacol. 90, 107152 https://doi.org/10.1016/J.INTIMP.2020.107152.

Maiese, K., 2016. Forkhead transcription factors: new considerations for alzheimer's disease and dementia. J. Transl. Sci. 2, 241. https://doi.org/10.15761/JTS.1000146.

Maiese, K., 2021. Targeting the core of neurodegeneration: FoxO, mTOR, and SIRT1. Neural Regen. Res. 16, 448. https://doi.org/10.4103/1673-5374.291382.

Mangiarini, L., Sathasivam, K., Seller, M., Cozens, B., Harper, A., Hetherington, C., Lawton, M., Trottier, Y., Lehrach, H., Davies, S.W., Bates, G.P., 1996. Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. Cell 87, 493–506. https://doi.org/10.1016/S0092-8674(00)81369-0.

Mani, S., 2015. Production of reactive oxygen species and its implication in human diseases. Free Radicals in Human Health and Disease. Springer, India, New Delhi, pp. 3–15. https://doi.org/10.1007/978-81-322-2035-0_1.

Martin, E., Heidari, R., Monnier, V., Tricoire, H., 2021. Genetic screen in adult Drosophila Reveals that dCBP depletion in glial cells mitigates huntington disease pathology through a foxo-dependent pathway. Int. J. Mol. Sci. 22 https://doi.org/10.3390/IJMS22083884. Page 3884 22, 3884. 2021.

Martins, R., Lithgow, G.J., Link, W., 2016. Long live FOXO: unraveling the role of FOXO proteins in aging and longevity. Aging Cell 15, 196–207. https://doi.org/10.1111/ACEL.12427.

Matsuoka, Y., Picciano, M., La Francois, J., Duff, K., 2001. Fibrillar β-amyloid evokes oxidative damage in a transgenic mouse model of Alzheimer's disease. Neuroscience 104, 609–613. https://doi.org/10.1016/S0306-4522(01)00115-4.

Meister, S., Storck, S.E., Hameister, E., Behl, C., Weggen, S., Clement, A.M., Pietrzik, C. U., 2015. Expression of the ALS-causing variant hSOD1G93A leads to an impaired integrity and altered regulation of claudin-5 expression in an in vitro blood–spinal cord barrier model. J. Cereb. Blood Flow Metab. 35, 1112. https://doi.org/10.1038/JCBFM.2015.57.

Murakami, K., Murata, N., Noda, Y., Tahara, S., Kaneko, T., Kinoshita, N., Hatsuta, H., Murayama, S., Barnham, K.J., Irie, K., Shirasawa, T., Shimizu, T., 2011. SOD1 (copper/zinc superoxide dismutase) deficiency drives amyloid β protein oligomerization and memory loss in mouse model of Alzheimer disease. J. Biol. Chem. 286, 44557–44568. https://doi.org/10.1074/jbc.M111.279208.

Mutlu, A.S., Gao, S.M., Zhang, H., Wang, M.C., 2020. Olfactory specificity regulates lipid metabolism through neuroendocrine signaling in Caenorhabditis elegans. Nat. Commun. 2020, 1–15. https://doi.org/10.1038/s41467-020-15296-8, 111 11.

Nagashima, T., Iino, Y., Tomioka, M., 2019. DAF-16/FOXO promotes taste avoidance learning independently of axonal insulin-like signaling. PLoS Genet. 15, e1008297 https://doi.org/10.1371/JOURNAL.PGEN.1008297.

Nemoto, S., 2002. Redox regulation of forkhead proteins through a p66shc-Dependent signaling pathway. Science 295 (80), 2450–2452. https://doi.org/10.1126/science.1069004.

Neri, C., 2012. Role and therapeutic potential of the pro-longevity factor FOXO and its regulators in neurodegenerative disease. Front. Pharmacol. 3, 15. https://doi.org/10.3389/FPHAR.2012.00015.

Oddo, S., Caccamo, A., Shepherd, J.D., Murphy, M.P., Golde, T.E., Kayed, R., Metherate, R., Mattson, M.P., Akbari, Y., LaFerla, F.M., 2003. Triple-transgenic model of Alzheimer's disease with plaques and tangles. Neuron 39, 409–421. https://doi.org/10.1016/S0896-6273(03)00434-3.

Ogunsuyi, O.B., Olagoke, O.C., Afolabi, B.A., Oboh, G., Ijomone, O.M., Barbosa, N.V., Rocha, J.B.Tda, 2021. Dietary Inclusions of Solanum vegetables Mitigate Aluminum-induced Redox and Inflammation-related Neurotoxicity in Drosophila melanogaster Model. https://doi.org/10.1080/1028415X.2021.1933331 https://doi.org/. doi: 10.1080/1028415X.2021.1933331.

Ojo, J.-O., Mouzon, B., Greenberg, M.B., Bachmeier, C., Mullan, M., Crawford, F., 2013. Repetitive mild traumatic brain injury augments tau pathology and glial activation in aged hTau mice. J. Neuropathol. Exp. Neurol. 72, 137–151. https://doi.org/10.1097/NEN.0b013e3182814cdf.

Paiva, F.A., Bonomo, L.de F., Boasquivis, P.F., de P.aula, I.T.B.R., Guerra, J.Fda C., Leal, W.M., Silva, M.E., Pedrosa, M.L., Oliveira, R.de P., 2015. Carqueja (Baccharis trimera) protects against Oxidative Stress and β-Amyloid-Induced Toxicity in Caenorhabditis elegans. Oxid. Med. Cell. Longev. 2015, 740162 https://doi.org/10.1155/2015/740162.

Park, D., Hahm, J.-H., Park, S., Ha, G., Chang, G.-E., Jeong, H., Kim, H., Kim, S., Cheong, E., Paik, Y.-K., 2017. A conserved neuronal DAF-16/FoxO plays an important role in conveying pheromone signals to elicit repulsion behavior in Caenorhabditis elegans. Sci. Rep. 7, 7260. https://doi.org/10.1038/s41598-017-07313-6.

Paroni, G., Seripa, D., Fontana, A., D'Onofrio, G., Gravina, C., Urbano, M., Cascavilla, L., Pellegrini, F., Greco, A., Pilotto, A., 2014. FOXO1 locus and acetylcholinesterase inhibitors in elderly patients with Alzheimer's disease. Clin. Interv. Aging 9, 1783. https://doi.org/10.2147/CIA.S64758.

Peng, C., Trojanowski, J.Q., Lee, V.M.-Y., 2020. Protein transmission in neurodegenerative disease. Nat. Rev. Neurol. 2020, 199–212. https://doi.org/10.1038/s41582-020-0333-7, 164 16.

Pino, E., Amamoto, R., Zheng, L., Cacquevel, M., Sarria, J.-C., Knott, G.W., Schneider, B. L., 2014. FOXO3 determines the accumulation of α-synuclein and controls the fate of dopaminergic neurons in the substantia nigra. Hum. Mol. Genet. 23, 1435–1452. https://doi.org/10.1093/HMG/DDT530.

Polter, A., Yang, S., Zmijewska, A.A., van Groen, T., Paik, J.-H., DePinho, R.A., Peng, S.L., Jope, R.S., Li, X., 2009. Forkhead Box, class O transcription factors in brain: regulation and behavioral manifestation. Biol. Psychiatry 65, 150–159. https://doi.org/10.1016/j.biopsych.2008.08.005.

Pratiwi, R., Nantasenamat, C., Ruankham, W., Suwanjang, W., Prachayasittikul, V., Prachayasittikul, S., Phopin, K., 2021. Mechanisms and neuroprotective activities of stigmasterol against oxidative stress-induced neuronal cell death via sirtuin family. Front. Nutr. 8, 648995 https://doi.org/10.3389/FNUT.2021.648995.

Qin, H., Zhang, H., Zhang, X., Zhang, S., Zhu, S., Wang, H., 2021. Resveratrol attenuates radiation enteritis through the SIRT1/FOXO3a and PI3K/AKT signaling pathways. Biochem. Biophys. Res. Commun. 554, 199–205. https://doi.org/10.1016/J.BBRC.2021.03.122.

Reiterer, M., Milton, S.L., 2020. Induction of foxo3a protects turtle neurons against oxidative stress. Comp. Biochem. Physiol., Part A Mol. Integr. Physiol. 243, 110671 https://doi.org/10.1016/J.CBPA.2020.110671.

Rena, G., Prescott, A.R., Guo, S., Cohen, P., Unterman, T.G., 2001. Roles of the forkhead in rhabdomyosarcoma (FKHR) phosphorylation sites in regulating 14-3-3 binding, transactivation and nuclear targetting. Biochem. J. 354, 605–612. https://doi.org/10.1042/0264-6021:3540605.

Romeo-Guitart, D., Marcos-DeJuana, C., Marmolejo-Martínez-Artesero, S., Navarro, X., Casas, C., 2020. Novel neuroprotective therapy with NeuroHeal by autophagy induction for damaged neonatal motoneurons. Theranostics 10, 5154–5168. https://doi.org/10.7150/THNO.43765.

Ruankham, W., Suwanjang, W., Wongchitrat, P., Prachayasittikul, V., Prachayasittikul, S., Phopin, K., 2019. Sesamin and Sesamol Attenuate H2O2-induced Oxidative Stress on Human Neuronal Cells Via the SIRT1-SIRT3-FOXO3a Signaling Pathway. https://doi.org/. doi: 10.1080/1028415X.2019.1596613 24, pp. 90–101. https://doi.org/10.1080/1028415X.2019.1596613.

Salih, D.A.M., Rashid, A.J., Colas, D., de la Torre-Ubieta, L., Zhu, R.P., Morgan, A.A., Santo, E.E., Ucar, D., Devarajan, K., Cole, C.J., Madison, D.V., Shamloo, M., Butte, A. J., Bonni, A., Josselyn, S.A., Brunet, A., 2012. FoxO6 regulates memory consolidation and synaptic function. Genes Dev. 26, 2780–2801. https://doi.org/10.1101/gad.208926.112.

Salimian, N., Peymani, M., Ghaedi, K., Nasr Esfahani, M.H., 2018. Modulation in miR-200a/SIRT1axis is associated with apoptosis in MPP+-induced SH-SY5Y cells. Gene 674, 25–30. https://doi.org/10.1016/J.GENE.2018.06.061.

Sang, T., Cao, Q., Wang, Y., Liu, F., Chen, S., 2014. Overexpression or silencing of FOXO3a affects proliferation of endothelial progenitor cells and expression of cell cycle regulatory proteins. PLoS One 9, e101703. https://doi.org/10.1371/journal.pone.0101703.

Sanphui, P., Das, A.K., Biswas, S.C., 2020. Forkhead Box O3a requires BAF57, a subunit of chromatin remodeler SWI/SNF complex for induction of p53 up-regulated modulator of apoptosis (Puma) in a model of Parkinson's disease. J. Neurochem. 154, 547–561. https://doi.org/10.1111/JNC.14969.

SantaCruz, K., 2005. Tau suppression in a neurodegenerative mouse model improves memory function. Science 309 (80), 476–481. https://doi.org/10.1126/science.1113694.

Santo, E.E., Paik, J., 2018. FOXO in neural cells and diseases of the nervous system. Curr. Top. Dev. Biol. 127, 105. https://doi.org/10.1016/BS.CTDB.2017.10.002.

Schäffner, I., Minakaki, G., Khan, M.A., Balta, E.-A., Schlötzer-Schrehardt, U., Schwarz, T.J., Beckervordersandforth, R., Winner, B., Webb, A.E., DePinho, R.A., Paik, J., Wurst, W., Klucken, J., Lie, D.C., 2018. FoxO function is essential for maintenance of autophagic flux and neuronal morphogenesis in adult neurogenesis. Neuron 99, 1188. https://doi.org/10.1016/J.NEURON.2018.08.017.

Schapira, A.H., 2008. Mitochondria in the aetiology and pathogenesis of Parkinson's disease. Lancet Neurol. 7, 97–109. https://doi.org/10.1016/S1474-4422(07)70327-7.

Schilling, G., 1999. Intranuclear inclusions and neuritic aggregates in transgenic mice expressing a mutant N-terminal fragment of huntingtin [published erratum appears in Hum Mol Genet 1999 May;8(5):943]. Hum. Mol. Genet. 8, 397–407. https://doi.org/10.1093/hmg/8.3.397.

Senchuk, M.M., Dues, D.J., Schaar, C.E., Johnson, B.K., Madaj, Z.B., Bowman, M.J., Winn, M.E., Van Raamsdonk, J.M., 2018. Activation of DAF-16/FOXO by reactive oxygen species contributes to longevity in long-lived mitochondrial mutants in Caenorhabditis elegans. PLoS Genet. 14, e1007268 https://doi.org/10.1371/JOURNAL.PGEN.1007268.

Setsuie, R., Wang, Y.-L., Mochizuki, H., Osaka, H., Hayakawa, H., Ichihara, N., Li, H., Furuta, A., Sano, Y., Sun, Y.-J., Kwon, J., Kabuta, T., Yoshimi, K., Aoki, S., Mizuno, Y., Noda, M., Wada, K., 2007. Dopaminergic neuronal loss in transgenic mice expressing the Parkinson's disease-associated UCH-L1 I93M mutant. Neurochem. Int. 50, 119–129. https://doi.org/10.1016/j.neuint.2006.07.015.

Shi, C., Viccaro, K., Lee, H., Shah, K., 2016. Cdk5–Foxo3 axis: initially neuroprotective, eventually neurodegenerative in Alzheimer's disease models. J. Cell. Sci. 129, 1815–1830. https://doi.org/10.1242/JCS.185009.

Simunkova, M., Alwasel, S.H., Alhazza, I.M., Jomova, K., Kollar, V., Rusko, M., Valko, M., 2019. Management of oxidative stress and other pathologies in Alzheimer's disease. Arch. Toxicol. 2019, 2491–2513. https://doi.org/10.1007/S00204-019-02538-Y, 939 93.

Singh, P., Hanson, P.S., Morris, C.M., 2017. SIRT1 ameliorates oxidative stress induced neural cell death and is down-regulated in Parkinson's disease. BMC Neurosci. 2017, 1–13. https://doi.org/10.1186/S12868-017-0364-1, 181 18.

Song, P., Zou, M.-H., 2015. Roles of reactive oxygen species in physiology and pathology. Atherosclerosis. John Wiley & Sons, Inc, Hoboken, NJ, pp. 379–392. https://doi.org/10.1002/9781118828533.ch30.

Sousa, N.A., Oliveira, G.A.L., Oliveira, A.Pde, Lopes, A.L.F., Iles, B., Nogueira, K.M., Araújo, T.S.L., Souza, L.K.M., Araújo, A.R., Ramos-Jesus, J., Plácido, A., Amaral, C., Campelo, Y.D.M., Barbosa, E.A., Portugal, C.C., Socodato, R., Lobo, A., Relvas, J., Bemquerer, M., Eaton, P., Leite, J.R.S.A., Medeiros, J.V.R., 2020. Novel ocellatin peptides mitigate LPS-induced ROS formation and NF-kB activation in microglia and hippocampal neurons. Sci. Reports 2020, 1–16. https://doi.org/10.1038/s41598-020-59665-1, 101 10.

Stamer, K., Vogel, R., Thies, E., Mandelkow, E., Mandelkow, E.-M., 2002. Tau blocks traffic of organelles, neurofilaments, and APP vesicles in neurons and enhances oxidative stress. J. Cell Biol. 156, 1051–1063. https://doi.org/10.1083/jcb.200108057.

Su, B., Liu, H., Wang, X., Chen, S.G., Siedlak, S.L., Kondo, E., Choi, R., Takeda, A., Castellani, R.J., Perry, G., Smith, M.A., Zhu, X., Lee, H., 2009. Ectopic localization of FOXO3a protein in Lewy bodies in Lewy body dementia and Parkinson's disease. Mol. Neurodegener. 4, 32. https://doi.org/10.1186/1750-1326-4-32.

Tas, D., Stickley, L., Miozzo, F., Koch, R., Loncle, N., Sabado, V., Gnägi, B., Nagoshi, E., 2018. Parallel roles of transcription factors dFOXO and FER2 in the development and maintenance of dopaminergic neurons. PLoS Genet. 14, e1007271 https://doi.org/10.1371/JOURNAL.PGEN.1007271.

Tia, N., Singh, A.K., Pandey, P., Azad, C.S., Chaudhary, P., Gambhir, I.S., 2018. Role of Forkhead Box O (FOXO) transcription factor in aging and diseases. Gene 648, 97–105. https://doi.org/10.1016/J.GENE.2018.01.051.

Tissenbaum, H.A., 2018. DAF-16: FOXO in the context of C. elegans. Curr. Top. Dev. Biol. 127, 1–21. https://doi.org/10.1016/BS.CTDB.2017.11.007.

Tousoulis, D., Andreou, I., Antoniades, C., Tentolouris, C., Stefanadis, C., 2008. Role of inflammation and oxidative stress in endothelial progenitor cell function and mobilization: therapeutic implications for cardiovascular diseases. Atherosclerosis 201.

Tran, H., 2002. DNA repair pathway stimulated by the forkhead transcription factor FOXO3a through the Gadd45 protein. Science 296 (80), 530–534. https://doi.org/10.1126/science.1068712.

Tripodi, F., Lombardi, L., Guzzetti, L., Panzeri, D., Milanesi, R., Leri, M., Bucciantini, M., Angeloni, C., Beghelli, D., Hrelia, S., Onorato, G., Schiavi, E.Di, Falletta, E., Nonnis, S., Tedeschi, G., Labra, M., Coccetti, P., 2020. Protective effect of Vigna unguiculata extract against aging and neurodegeneration. Aging (Albany NY) 12, 19785. https://doi.org/10.18632/AGING.104069.

Trist, B.G., Hare, D.J., Double, K.L., 2019. Oxidative stress in the aging substantia nigra and the etiology of Parkinson's disease. Aging Cell 18, e13031. https://doi.org/10.1111/ACEL.13031.

Urbánek, P., Klotz, L., 2017. Posttranscriptional regulation of FOXO expression: microRNAs and beyond. Br. J. Pharmacol. 174, 1514. https://doi.org/10.1111/BPH.13471.

Urbich, C., Dimmeler, S., 2004. Endothelial progenitor cells: characterization and role in vascular biology. Circ. Res. 95.

Vidal, R.L., Figueroa, A., Court, F.A., Thielen, P., Molina, C., Wirth, C., Caballero, B., Kiffin, R., Segura-Aguilar, J., Cuervo, A.M., Glimcher, L.H., Hetz, C., 2012. Targeting the UPR transcription factor XBP1 protects against Huntington's disease through the regulation of FoxO1 and autophagy. Hum. Mol. Genet. 21, 2245. https://doi.org/10.1093/HMG/DDS040.

Voisin, J., Farina, F., Naphade, S., Fontaine, M., Tshilenge, K.-T., Aguirre, C.G., Lopez-Ramirez, A., Dancourt, J., Ginisty, A., Nair, S.S., Madushani, K.L., Zhang, N., Lejeune, F.-X., Verny, M., Campisi, J., Ellerby, L.M., Neri, C., 2020. FOXO3 targets are reprogrammed as Huntington's disease neural cells and striatal neurons face senescence with p16INK4a increase. Aging Cell 19, e13226. https://doi.org/10.1111/ACEL.13226.

Wang, F., Wang, Y., Cao, Q., Zhang, J., Huang, L., 2013. Hydrogen peroxide induced impairment of endothelial progenitor cell viability is mediated through a FoxO3a dependant mechanism. Microvasc. Res.

Wang, X., Wang, W., Li, L., Perry, G., Lee, H., Zhu, X., 2014a. Oxidative stress and mitochondrial dysfunction in Alzheimer's disease. Biochim. Biophys. Acta - Mol. Basis Dis. 1842, 1240–1247. https://doi.org/10.1016/j.bbadis.2013.10.015.

Wang, Y., Zhou, Y., Graves, D.T., 2014b. FOXO transcription factors: their clinical significance and regulation. Biomed Res. Int. 2014 https://doi.org/10.1155/2014/925350.

Wang, Y.-Q., Cao, Q., Wang, F., Huang, L.-Y., Sang, T.-T., Liu, F., Chen, S.-Y., 2015. SIRT1 protects against oxidative stress-induced endothelial progenitor cells apoptosis by inhibiting FOXO3a via FOXO3a ubiquitination and degradation. J. Cell. Physiol. 230, 2098–2107. https://doi.org/10.1002/JCP.24938.

Wang, Z., Yu, T., Huang, P., 2016. Post-translational modifications of FOXO family proteins (Review). Mol. Med. Rep. 14, 4931–4941. https://doi.org/10.3892/MMR.2016.5867.

Wang, B., Li, Y., Wang, H., Zhao, J., Zhao, Y., Liu, Z., Ma, H., 2020a. FOXO3a is stabilized by USP18-mediated de-ISGylation and inhibits TGF-β1-induced fibronectin expression. J. Investig. Med. 68, 786–791. https://doi.org/10.1136/JIM-2019-001145.

Wang, Y., Lin, Y., Wang, L., Zhan, H., Luo, X., Zeng, Y., Wu, W., Zhang, X., Wang, F., 2020b. TREM2 ameliorates neuroinflammatory response and cognitive impairment via PI3K/AKT/FoxO3a signaling pathway in Alzheimer's disease mice. Aging (Albany. NY). 12, 20862–20879. https://doi.org/10.18632/AGING.104104.

Wang, J., Liu, B., Xu, Y., Yang, M., Wang, C., Song, M., Liu, J., Wang, W., You, J., Sun, F., Wang, D., Liu, D., Yan, H., 2021. Activation of CREB-mediated autophagy by thioperamide ameliorates β-amyloid pathology and cognition in Alzheimer's disease. Aging Cell 20, e13333. https://doi.org/10.1111/ACEL.13333.

Weller, J., Budson, A., 2018. Current understanding of Alzheimer's disease diagnosis and treatment. F1000Research 7. https://doi.org/10.12688/F1000RESEARCH.14506.1.

Werner, N., Kosiol, S., Schiegl, T., Ahlers, P., Walenta, K., 2005. Circulating endothelial progenitor cells and cardiovascular outcomes. N. Engl. J. Med. 353.

Wheeler, V.C., 2000. Long glutamine tracts cause nuclear localization of a novel form of huntingtin in medium spiny striatal neurons in HdhQ92 and HdhQ111 knock-in mice. Hum. Mol. Genet. 9, 503–513. https://doi.org/10.1093/hmg/9.4.503.

Willcox, B.J., Donlon, T.A., He, Q., Chen, R., Grove, J.S., Yano, K., Masaki, K.H., Willcox, D.C., Rodriguez, B., Curb, J.D., 2008. FOXO3A genotype is strongly associated with human longevity. Proc. Natl. Acad. Sci. 105, 13987–13992. https://doi.org/10.1073/pnas.0801030105.

Wong, H.-K.A., Veremeyko, T., Patel, N., Lemere, C.A., Walsh, D.M., Esau, C., Vanderburg, C., Krichevsky, A.M., 2013. De-repression of FOXO3a death axis by microRNA-132 and -212 causes neuronal apoptosis in Alzheimer's disease. Hum. Mol. Genet. 22, 3077–3092. https://doi.org/10.1093/HMG/DDT164.

Wong, S.Q., Pontifex, M.G., Phelan, M.M., Pidathala, C., Kraemer, B.C., Barclay, J.W., Berry, N.G., O'Neill, P.M., Burgoyne, R.D., Morgan, A., 2018. α-Methyl-α-phenylsuccinimide ameliorates neurodegeneration in a C. Elegans model of TDP-43 proteinopathy. Neurobiol. Dis. 118, 40–54. https://doi.org/10.1016/J.NBD.2018.06.013.

Wongchai, K., Schlotterer, A., Lin, J., Humpert, P.M., Klein, T., Hammes, H.-P., Morcos, M., 2015. Protective effects of Liraglutide and Linagliptin in C. Elegans as a new model for glucose-induced neurodegeneration. Horm. Metab. Res. 48, 70–75. https://doi.org/10.1055/S-0035-1549876.

Wu, Q., Yuan, X., Bai, J., Han, R., Li, Z., Zhang, H., Xiu, R., 2019. MicroRNA-181a protects against pericyte apoptosis via directly targeting FOXO1: implication for ameliorated cognitive deficits in APP/PS1 mice. Aging (Albany. NY). 11, 6120–6133. https://doi.org/10.18632/AGING.102171.

Wu, B.-W., Guo, J.-D., Wu, M.-S., Liu, Y., Lu, M., Zhou, Y.-H., Han, H.-W., 2020. Osteoblast-derived lipocalin-2 Regulated by miRNA-96-5p/Foxo1 Advances the Progression of Alzheimer's Disease. https://doi.org/10.2217/epi-2019-0215 12, pp. 1501–1513. https://doi.org/10.2217/EPI-2019-0215.

Xia, E., Xu, F., Hu, C., Kumal, J.P.P., Tang, X., Mao, D., Li, Y., Wu, D., Zhang, R., Wu, S., Sun, L., 2019. Young blood rescues the cognition of alzheimer's model mice by restoring the hippocampal cholinergic circuit. Neuroscience 417, 57–69. https://doi.org/10.1016/J.NEUROSCIENCE.2019.08.010.

Xia, S., Wen, X., Fan, X., Chen, X., Wei, Z., Li, Q., Sun, L., 2020. Wnt2 overexpression protects against PINK1 mutant-induced mitochondrial dysfunction and oxidative stress. Mol. Med. Rep. 21, 2633–2641. https://doi.org/10.3892/MMR.2020.11066.

Xu, S., Stern, M., McNew, J.A., 2017a. Beneficial effects of rapamycin in a Drosophila model for hereditary spastic paraplegia. J. Cell. Sci. 130, 453. https://doi.org/10.1242/JCS.196741.

Xu, X., Zhang, L., Ye, X., Hao, Q., Zhang, T., Cui, G., Yu, M., 2017b. Nrf2/ARE pathway inhibits ROS-induced NLRP3 inflammasome activation in BV2 cells after cerebral ischemia reperfusion. Inflamm. Res. 2017, 57–65. https://doi.org/10.1007/S00011-017-1095-6, 671 67.

Yang, C.-C., Hsiao, L.-D., Yang, C.-M., 2020. Galangin Inhibits LPS-Induced MMP-9 Expression via Suppressing Protein Kinase-Dependent AP-1 and FoxO1 Activation in Rat Brain Astrocytes . J. Inflamm. Res. 13, 945–960. https://doi.org/10.2147/JIR.S276925.

Yang, W., Wang, Y., Hao, Y., Wang, Z., Liu, J., Wang, J., 2021. Piceatannol alleviate ROS-mediated PC-12 cells damage and mitochondrial dysfunction through SIRT3/FOXO3a signaling pathway. J. Food Biochem., e13820 https://doi.org/10.1111/JFBC.13820.

Yen, K., Mehta, H.H., Kim, S.J., Lue, Y.H., Hoang, J., Guerrero, N., Port, J., Bi, Q., Navarrete, G., Brandhorst, S., Lewis, K.N., Wan, J., Swerdloff, R., Mattison, J.A., Buffenstein, R., Breton, C.V., Wang, C., Long, V., Atzmon, G., Wallace, D., Barzilai, N., Cohen, P., 2020. The mitochondrial derived peptide humanin is a regulator of lifespan and healthspan. Aging (Albany. NY). 12, 11185–11199. https://doi.org/10.18632/AGING.103534.

Yoshiyama, Y., Higuchi, M., Zhang, B., Huang, S.-M., Iwata, N., Saido, T.C., Maeda, J., Suhara, T., Trojanowski, J.Q., Lee, V.M.-Y., 2007. Synapse loss and microglial activation precede tangles in a P301S tauopathy mouse model. Neuron 53, 337–351. https://doi.org/10.1016/j.neuron.2007.01.010.

Yossef, R.R., Al-Yamany, M.F., Saad, M.A., El-Sahar, A.E., 2020. Neuroprotective effects of vildagliptin on drug induced Alzheimer's disease in rats with metabolic syndrome: role of hippocampal klotho and AKT signaling pathways. Eur. J. Pharmacol. 889, 173612 https://doi.org/10.1016/J.EJPHAR.2020.173612.

You, M., Miao, Z., Tian, J., Hu, F., 2019. Trans -10-hydroxy-2-decenoic acid protects against LPS-induced neuroinflammation through FOXO1-mediated activation of autophagy. Eur. J. Nutr. 2019, 2875–2892. https://doi.org/10.1007/S00394-019-02128-9, 597 59.

Yusuf, I.O., Chen, H.-M., Cheng, P.-H., Chang, C.-Y., Tsai, S.-J., Chuang, J.-I., Wu, C.-C., Huang, B.-M., Sun, H.S., Chen, C.-M., Yang, S.-H., 2021. Fibroblast growth factor 9 stimulates neuronal length through NF-kB signaling in striatal cell huntington's disease models. Mol. Neurobiol. 2021, 2396–2406. https://doi.org/10.1007/S12035-020-02220-W, 585 58.

Zhang, T., Baldie, G., Periz, G., Wang, J., 2014. RNA-processing protein TDP-43 regulates FOXO-Dependent protein quality control in stress response. PLoS Genet. 10, e1004693 https://doi.org/10.1371/journal.pgen.1004693.

Zhang, J., Shi, R., Li, H., Xiang, Y., Xiao, L., Hu, M., Ma, F., Ma, C.W., Huang, Z., 2016. Antioxidant and neuroprotective effects of Dictyophora indusiata polysaccharide in Caenorhabditis elegans. J. Ethnopharmacol. 192, 413–422. https://doi.org/10.1016/J.JEP.2016.09.031.

Zhang, C., Tan, Z., Xie, Y., Zhao, Y., Huang, T.Y., Lu, Z., Luo, H., Can, D., Xu, H., Zhang, Y., Zhang, X., 2019. Appoptosin mediates lesions induced by oxidative stress through the JNK-FoxO1 pathway. Front. Aging Neurosci. 0, 243. https://doi.org/10.3389/FNAGI.2019.00243.

Zhang, W., Bai, S., Yang, J., Zhang, Y., Liu, Y., Nie, J., Meng, D., Shi, R., Yao, Z., Wang, M., Wang, H., Li, C., 2020. FoxO1 overexpression reduces Aβ production and tau phosphorylation in vitro. Neurosci. Lett. 738, 135322 https://doi.org/10.1016/J.NEULET.2020.135322.

Zhao, Y., Wang, Y., Zhu, W.-G., 2011. Applications of post-translational modifications of FoxO family proteins in biological functions. J. Mol. Cell Biol. 3, 276–282. https://doi.org/10.1093/JMCB/MJR013.

Zhao, Y., Zhao, R., Wu, J., Wang, Q., Pang, K., Shi, Q., Gao, Q., Hu, Y., Dong, X., Zhang, J., Sun, J., 2018. Melatonin protects against Aβ-induced neurotoxicity in primary neurons via miR-132/PTEN/AKT/FOXO3a pathway. BioFactors 44, 609–618. https://doi.org/10.1002/BIOF.1411.

Zhao, Y., Zhao, B., 2013. Oxidative stress and the pathogenesis of Alzheimer's disease. Oxid. Med. Cell. Longev. 2013, 1–10. https://doi.org/10.1155/2013/316523.

Zheng, H., Jiang, M., Trumbauer, M.E., Sirinathsinghji, D.J.S., Hopkins, R., Smith, D.W., Heavens, R.P., Dawson, G.R., Boyce, S., Conner, M.W., Stevens, K.A., Slunt, H.H., Sisodia, S.S., Chen, H.Y., Van der Ploeg, L.H.T., 1995. β-amyloid precursor protein-

deficient mice show reactive gliosis and decreased locomotor activity. Cell 81, 525–531. https://doi.org/10.1016/0092-8674(95)90073-X.

Zhou, J., Li, H., Li, X., Zhang, G., Niu, Y., Yuan, Z., Herrup, K., Zhang, Y.-W., Bu, G., Xu, H., Zhang, J., 2015. The roles of Cdk5-Mediated subcellular localization of FOXO1 in neuronal death. J. Neurosci. 35, 2624–2635. https://doi.org/10.1523/JNEUROSCI.3051-14.2015.

Zou, D., Mou, Z., Wu, W., Liu, H., 2021. TRIM33 protects osteoblasts from oxidative stress-induced apoptosis in osteoporosis by inhibiting FOXO3a ubiquitylation and degradation. Aging Cell, e13367. https://doi.org/10.1111/ACEL.13367.

2021 Fourth International Conference on Computational Intelligence and Communication Technologies (CCICT) | 978-1-6654-2392-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/CCICT53244.2021.00025

FPGA based design of multifunction ALU

Anup Kumar[1], Manish Kumar[2], Gouri Shankar Jha[3], Kriti Suneja[4]
Dept. of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
anup718199@gmail.com[1], manish.kv17@gmail.com[2], gourishankarjha1011@gmail.com[3], kritisuneja@dtu.ac.in[4]

***Abstract*— With the growing need of fast arithmetic calculations, Field Programmable Gate Arrays (FPGAs) have shown the promising results. To take this idea further, we have implemented five functions which are of immense use in digital signal processing, communication, encryption, etc. applications. The simulations have been performed in Xilinx Vivado and the target device for synthesis results belongs to Kintex FPGA family. The simulation results are in good agreement with the theory and synthesis results show the device utilization in the form of Look Up tables, an indicator of area used. The comparison among different implemented functions shows that while average requires least number of LUTs, circular convolution requires most.**

Keywords*— *Field Programmable Gate Arrays, Arithmetic Logic Unit, Synthesis, Convolution

I. INTRODUCTION

An Arithmetic Logic Unit (ALU) consists of multiple mathematical and logical operations in it. Previously, many ALUs were designed with various operations which is mentioned in the Table I, with their authors and target FPGA Device on which operations were implemented.

This research paper focuses on ALU with operations which are specifically useful in digital signal processing applications. These operations are specified in detail in Section II of this research paper.

In order to obtain results of different mathematical computational data which are used in our engineering field, which are also complex in nature, time consuming and not too error free when calculated manually, we have proposed Field Programmable Gate Arrays (FPGAs) based design of ALU. To overcome the risk of errors in it and to calculate and verify these computational functions in a less period of time, we have implemented these functions using Verilog Hardware Description Language in Xilinx EDA tool. Further, we combine all these functions to form an ALU so that a single designed chip can perform all the individually implemented functions.

The vast amount of literature in designing of FPGA based ALUs in the previous years clearly depicts their importance among electronic engineers. In [1], authors have implemented the Sine and Cosine function in FPGA using Cordic Algorithm for determining Satellite Attitude in their proposed system. In [2], sine and cosine functions have been implemented on Virtex device. The rest of the available work in literature has been summarized in Table I.

Table I: ALU Operations Implemented

Author and Reference	Operations Implemented	Target FPGA Device
Shoaib Bhuria, P.Muralidhar [1]	Sine and Cosine	Altera's Quartus II 9.0
Shrugal Varde et al [2]	Hyperbolic Tan	Virtex5
Shikha Khurana, Kanika Kaur [3]	Adder, subtractor, multiplier, divisor, Logical Operations: AND, OR, NOR, Buffer, NAND, XOR, Inverter	Spartan
Nidhi Khanna, Dr. D.K.Mishra [4]	Addition, Subtraction, Increment, Decrement, Clear, NOR, XOR, XNOR, AND, NAND, OR, Complement, Shift left, Shift Right	Artix 7
Mrs. Roopa Kulkarni, Dr. S.Y. Kulkarni [5]	Increment, Decrement, Arithmetic shift left, Addition, Subtract, Add with carry, Subtract with borrow, Clear, Out B, Complement B, Out A, Complement B, Logical shift left, AND operation	SPARTAN6
S.M. Swamynathan, V.Banumathi [6]	Add, Subtract, division, multiplication, AND, OR, XOR, XNOR, NAND, inverter	---
Chunyang Feng, Liang Yang [7]	Addition/ Subtraction, Multiplier, Unified FP	Virtex 6
Debarshi Datta et al [8]	Multibit LFSR	Spartan 6
Yiyu Tan, Toshiyuki Imamura [9]	Matrix Multiplier	DE5a-NET
Stefan Lachowiczm, Hans-Jörg Pfleiderer [10]	Square Root, Look-up tables, CORDIC core, Series expansion, Newton-Raphson iteration	Virtex 2
Amit Kumar Panda et al [11]	8, 16 and 32 Bit LFSR	Spartan 3S

978-1-6654-2392-2/21/$31.00 ©2021 IEEE
DOI 10.1109/CCICT53244.2021.00025

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:12:26 UTC from IEEE Xplore. Restrictions apply.

The proposed ALU which we have designed can implement signals based computations like correlation, convolution etc., and also the functions given below.

1. Matrix Multiplication [9]
2. Correlation
3. Circular Convolution [12]
4. Coordinate System
5. Average
6. LFSR [8,11]

All these functions are first implemented individually, then they are merged in a way, so, that a single FPGA Chip can be used for performing all these functions.

For implementing these functions, we have used Verilog HDL in Xilinx Vivado Ver.2016.3 for writing and simulating the codes. The target FPGA device is XCKU035-FFVA1156-3-E, which belongs to the family Kintex Ultrascale FPGA.

The rest of the paper is organized as follows: Section II describes the functions which have been implemented and their simulation results. Section III describes the synthesis results and finally Section IV concludes the paper.

II. IMPLEMENTED OPERATIONS

A. MATRIX MULTIPLICATION

Matrix Multiplication is a mathematical function. In which two matrices are multiplied which can be in form of m*n or n*m, where m is the number of rows and n is the number of columns.

In this implementation we have used the data inside the matrix in boolean form and calculated it in boolean.

In this system the number of columns in X_{matrix} is equal to number of rows in Y_{matrix}. The Final Result is named as Matrix Product and it is denoted as XY.

For E.g.; we have taken two matrices of 2x2 and multiplies them,

$$\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix} \quad x \quad \begin{bmatrix}5 & 6\\7 & 8\end{bmatrix} \quad = \quad \begin{bmatrix}19 & 22\\43 & 50\end{bmatrix}$$

Matrix X Matrix Y Matrix XY

Input and Output of the matrix multiplication operation is shown in Fig 1 and Fig 2.

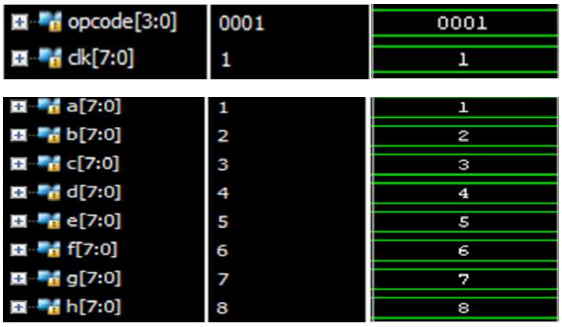


Fig 1. Matrix X and Matrix Y Inputs values

The Data taken in simulations for matrices X and Y is entered as input in the matrix multiplier. Fig.1 shows that the input data, where a, b, c, d are the components of matrix X and e, f, g, h are components of matrix Y. Fig.2 shows the result of the multiplication operation and the resultant matrix is named as Matrix XY here.

B. CORRELATION

Correlation function is used in signal processing to measure the degree of resemblance between two signals. In this system, we have implemented all three types of correlations as shown below:

1. Cross Correlation: is a measure of similarity of two series as a function of the displacement of one relative to the other.
2. Auto Correlation: In Auto Correlation we find the similarity between the signals and copy of same signal with delay, the delayed copy of signal is known as function of delay.
3. Circular Correlation: it calculates the similarity between the signals.

The simulation results of all the three types of correlations have been shown in Fig. 3.

Fig. 3 (a) shows the input which was entered for performing the correlation function.

Fig. 3 (b), 3(c) and 3(d) show the result of the cross correlation, auto correlation and circular correlation respectively which is performed by the ALU on the entered input data.

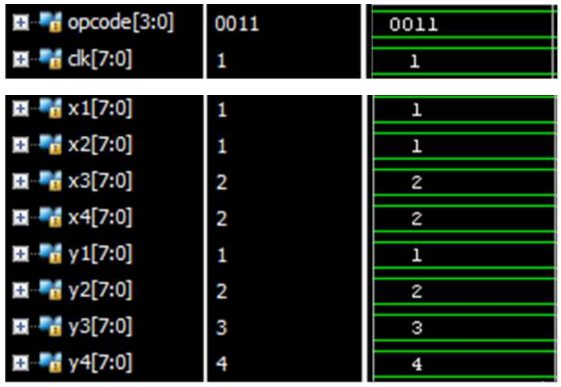


Fig. 3 (a) Input of Correlation function

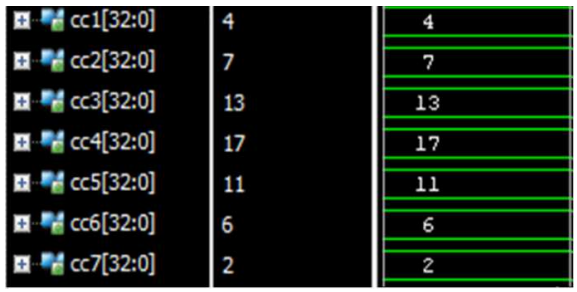


Fig. 3 (b) Cross correlation output

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:12:26 UTC from IEEE Xplore. Restrictions apply.

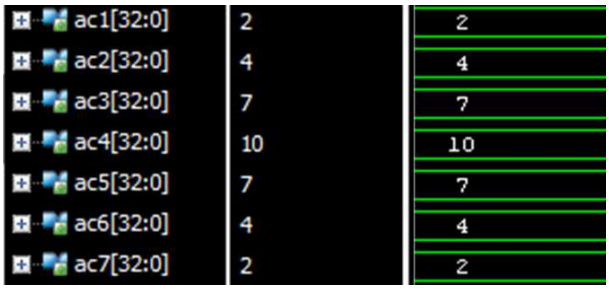


Fig. 3 (c) Auto correlation output

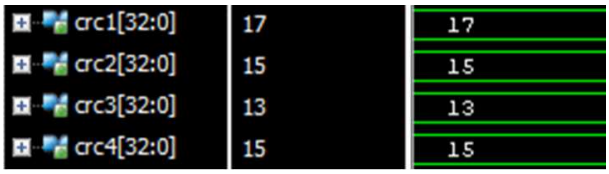


Fig. 3 (d) Circular correlation output

C. CIRCULAR CONVOLUTION

Circular convolution is also known as cyclic convolution, it is a special scenario of periodic convolution.

We have taken six values of X (x_{00}, x_{11}, x_{22}, x_{33}, x_{44}, x_{55}) and six values of Y (y_{00}, y_{11}, y_{22}, y_{33}, y_{44}, y_{55}) and use in the function of circular convolution.

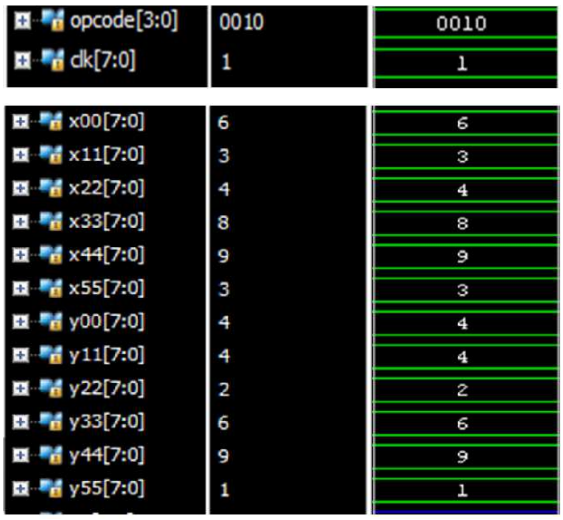


Fig 4. Input: X and Y inputs

Figure 4 shows the input values entered for the operation of circular convolution, where x00 to x55 are the values of X and y00 to y55 are values of Y.

Figure 5 shows the result of circular convolution operation, where z00 to z55 show the results of the operation.

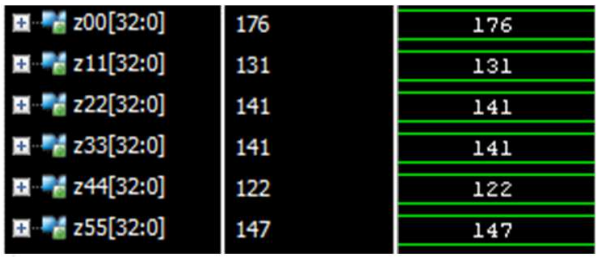


Fig 5. Circular Convolution Output

D. COORDINATE SYSTEM

Coordinate system is used to identify the location of a point, object or geometric element in a Euclidian Space.

In this proposed system we have implemented cylindrical and spherical coordinates.

In cylindrical coordinate a point is represented by ρ, ϕ, z and their value is calculated by following equations:

$$\rho = \sqrt{x^2 + y^2} \quad (1)$$

$$\phi = \tan^{-1}\frac{y}{x} \quad (2)$$

$$z = z \quad (3)$$

In spherical coordinates, a point is represented by r, θ, ϕ and their values are calculated by following equations:

$$r = \sqrt{x^2 + y^2 + z^2} \quad (4)$$

$$\theta = \tan^{-1}\frac{\sqrt{x^2+y^2}}{z} \quad (5)$$

$$\phi = \tan^{-1}\frac{y}{x} \quad (6)$$

Figure 6 and 7 give the corresponding simulation results.

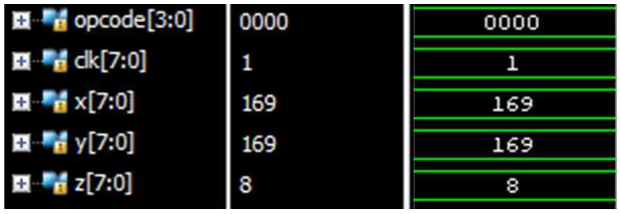


Fig 6. Input x and y and z

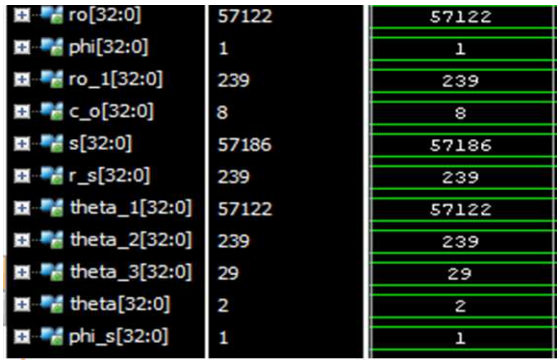


Fig 7. Calculated Results

E. AVERAGE

For the implementation of average function, in this n number are given each no. is expressed by I (where i = 1,2,…….n), the arithmetic mean is the sum of the value and divided by n as shown in (7).

$$\frac{1}{n}\sum_{i=1}^{n} a_i = \frac{a_1+a_2+\cdots+a_n}{n} \quad (7)$$

The Input and outputs results of the average operation are shown in Fig 8 & Fig 9.

Fig.8 shows the 13 entries (input) for finding the result of average and Fig. 9 shows the result of average operation of these 13 entries (input) for finding the result.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:12:26 UTC from IEEE Xplore. Restrictions apply.

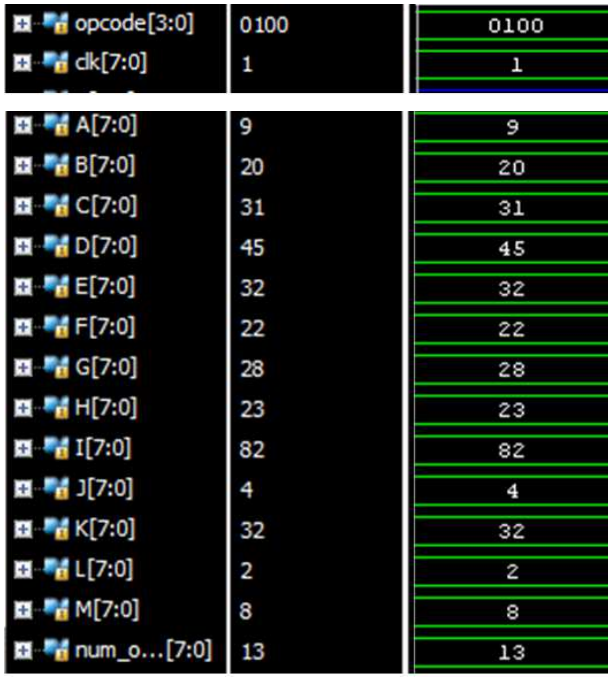


Fig 8. Input: Entered 13 random entries

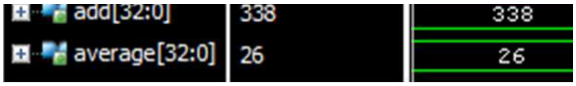


Fig 9. Output: average of 13 random entries

F. LFSR

LFSR [11] is a shift register in which input bit is linear function of its previous state. Commonly used linear function of single bit is XOR. LFSR is shift register whose input is controlled by XOR of some of the bits of the overall shift register value. The seed is provided to LFSR to start its initial cycle. LFSR is designed in such a manner so that the feedback function can generate a set of random numbers and a very long cycle.

We have designed a 8-bit LFSR and its output is shown in Fig 10.



Fig 10. Output of LFSR

III. PROPOSED SYSTEM

The functions specified in Section II of implemented operations are merged to form an ALU, so that a single FPGA board will be able to perform all these operations.

After merging all functions in a single code we have used designated op-codes to perform a specific operation. Once a user enters the variable, user needs to select the operation which is specified in op-code. After, entering that op-code system will perform that task.

These op-codes are given in Table II. For e.g.; if we select the op-code 0010 as specified in the table, and enters the input the system will move to do the circular convolution and gives the related result.

A. OP-CODES FOR DIFFERENT OPERATIONS

We have used 4-bit op-code, total 16 op-code can be implemented. Here, we have used 6 op-code and remaining are spare and can be used further. All op-code with their respective operation are shown in below Table II.

Table II: Op-code and Operations

S.No.	op-code	Operation
1.	0000	Coordinate system
2.	0001	Matrix Multiplication
3.	0010	Circular convolution
4.	0011	Correlation
5.	0100	Average
6.	0101	LFSR

B. LUT and POWER

We have checked the used Look Up Tables (LUTs) as per different operations and it is mentioned in below Table III.

Table III: LUTs & Power consumption

S.No.	Operations	LUT	Power
1.	Matrix Multiplication	283	20.478W
2.	Correlation	1298	182.229W
3.	Circular Convolution	1374	71.691W
4.	Coordinate System	195	45.606W
5.	Average	182	24.036W

Also, we have checked the LUTs which are used in when these all operations are merged together and form an ALU. The LUTs used and power consumption is shown in Table IV.

Table IV: LUT & Power used by ALU.

S.No.	Operations	LUT	Power
1.	ALU	1471	229.881W

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:12:26 UTC from IEEE Xplore. Restrictions apply.

IV. CONCLUSION

The proposed system where we have merged the six functions, which are matrix multiplication, correlation, convolution, coordinate system, average and LFSR have been implemented, and merged in an ALU and its power consumption and LUTs used are observed. It is to be noted here that the functions chosen are highly useful in signal processing and communication applications. Though these functions have been implemented on FPGA before, as available in literature, but a unified system to incorporate all these functions and offering an application specific ALU was the goal of this work which has been achieved to a large extent. Also, we have checked that how much power and LUTs are consumed by the individual functions. The results, both simulation and synthesis, are in good agreement with theory. We further propose to use this unified system in signal processing and communication applications.

REFERENCES

[1] S. Bhuria and P. Muralidhar, "FPGA implementation of sine and cosine value generators using Cordic Algorithm for Satellite Attitude Determination and calculators," 2010 International Conference on Power, Control and Embedded Systems, 2010, pp. 1-5, doi: 10.1109/ICPCES.2010.5698645.

[2] S. Varde, Dr. N. Sarwade and R. Upadhyay, "Hardware Implementation Of Hyperbolic Tan Using Cordic On FPGA," International Journal of Engineering Research and Applications (IJERA) , 2013,Vol. 3(2), pp.696-699.

[3] S. Khurana and K. Kaur, "IMPLEMENTATION OF ALU USING FPGA" International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), 2012, Vol. 1(2).

[4] N. Khanna and Dr. D.K.Mishra, "Clock Gated 16-Bits ALU Design & Implementation on FPGA" 2018 4th International Conference for Convergence in Technology (I2CT) SDMIT Ujire, Mangalore, India. Oct 27-28, 2018.

[5] Mrs. R. Kulkarni and Dr. S.Y.Kulkarni, "Power Analysis and Comparison of Clock Gated Techniques Implemented on a 16-bit ALU" Proceedings of International Conference on Circuits, Communication, Control and Computing (I4C 2014) MSRIT, BANGALORE, India, Nov. 21-22, 2014.

[6] S. M. Swamynathan and V. Banumathi, "Design and analysis of FPGA based 32 bit ALU using reversible gates," 2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE), 2017, pp. 1-4, doi: 10.1109/ICEICE.2017.8191959.

[7] C. Feng and L. Yang, "Design and evaluation of a novel reconfigurable ALU based on FPGA," Proceedings 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC), 2013, pp. 2286-2290, doi: 10.1109/MEC.2013.6885424.

[8] D. Datta, B. Datta and H. S. Dutta, "Design and implementation of multibit LFSR on FPGA to generate pseudorandom sequence number," 2017 Devices for Integrated Circuit (DevIC), 2017, pp. 346-349, doi: 10.1109/DEVIC.2017.8073966.

[9] Y. Tan and T. Imamura, "An energy-efficient FPGA-based matrix multiplier," 2017 24th IEEE International Conference on Electronics, Circuits and Systems (ICECS), 2017, pp. 514-517, doi: 10.1109/ICECS.2017.8292049.

[10] S. Lachowicz and H. Pfleiderer, "Fast Evaluation of the Square Root and Other Nonlinear Functions in FPGA," 4th IEEE International Symposium on Electronic Design, Test and Applications (delta 2008), 2008, pp. 474-477, doi: 10.1109/DELTA.2008.119.

[11] A. K. Panda, P. Rajput and B. Shukla, "FPGA Implementation of 8, 16 and 32 Bit LFSR with Maximum Length Feedback Polynomial Using VHDL," 2012 International Conference on Communication Systems and Network Technologies, 2012, pp. 769-773, doi: 10.1109/CSNT.2012.168.

[12] A. Burud and P. Bhaskar, "Design and Implementation of FPGA Based 32 Bit Floating Point Processor for DSP Application," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018, pp. 1-5, doi: 10.1109/ICCUBEA.2018.8697775.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:12:26 UTC from IEEE Xplore. Restrictions apply.

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

Free Space Optical Wireless Communication using Lasers Array for Data Transmission

Durgesh
Department of Electronics & Communication Engineering
Delhi Technological University
Delhi, India
durgeshlodhi327275@gmail.com

Deepak Ghanshala
Department of Electronics & Communication Engineering
Delhi Technological University
Delhi, India
deepakdaring90@gmail.com

Nachiketa Panwar
Department of Electronics & Communication Engineering
Delhi Technological University
Delhi, India
navneetsharma241@gmail.com

Piyush Tewari
Department of Electronics & Communication Engineering
Delhi Technological University
Delhi, India
piyushtewari@dtu.ac.in

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498473

***Abstract*—Free-space optical (FSO) communication using Light Fidelity (Li-Fi) is a booming field with several advantages over optical fiber communication such as low cost, lower losses, lower complexity, and higher data rate. The Data transmission with the help of LASERs through a transmitter suffers with various interruptions along the channel such as obstacles, rain, wind, and fog. In this work, we have proposed a Li-Fi technology-based reliable FSO wireless communication system model using LASERs. The suggested system is implemented using an array of LASERs for uninterrupted high-speed communication. Converging lens is used just before the receiver to converge the multiple LASER beams transmitted by the array of LASERs to the receiver. The converging lens also increases the intensity of the LASERs beams at the receiver side. In this work, first we investigate the performance of the proposed system in the OptiSystem tool in terms of quality factor (Q-factor), bit error rate (BER), received power, Eye diagram with respect to the link distance. We also present a hardware prototype for the proposed communication system and analyze its performance.**

Keywords— OptiSystem 7, Free Space Optical, FSO, Li-Fi, Bit Error Rate, BER, MISO, Array of LASERs, Attenuation in Different Weather Conditions, Rain, Haze, Fog

I. INTRODUCTION

With the rapid proliferation of modern wireless communication systems both in professional and personal life with the emergence of 5G communication, virtual reality (VR), internet of things (IoT), augmented reality (AR), xK Television, machine type communication (MTC), etc. the need for wireless communication resources is growing at an unprecedented rate [1]. To support these high data rates modern communication systems very high bandwidth resource is required which is difficult to be achieved with the existing radio frequency (RF) communication systems. As an effort to propose an effective solution to this situation, the scientific community and academia have recently turned their attention towards mm-wave and THz frequency communication systems. FSO communication is a promising communication technology to overcome these challenges and to establish a reliable, very-high-speed, and efficient communication system.

In [2] it has been shown that the RF-Microwave radiation is harmful to humans and birds as large high power Antennas produce harmful radiations. This situation can be avoided by using LASERs beams for data transmission, as used in Li-Fi based FSO wireless communication system. FSO wireless communication system has several other advantages over the RF-Microwave communication system, such as high data rate, higher security [5], and lower complexity.

Further, the data rate can be improved manifold by using the multiple-input multiple-output (MIMO) method in the FSO communication system [3]. In [3] a wavelength division multiplexing (WDM) FSO communication system transmits hybrid high-speed signals combined with different wavelengths using a single beam to achieve a high-speed hybrid signal [3]. In [4] data rate transmission up to 25Gbps was achieved using the MIMO technique with DEMUX in a WDM-FSO system. In this work, a 2D-Photo detector array was used for receiving multiple signals simultaneously at the receiver.

The FSO communication system also suffers from several challenges such as interruption to the transmitted signal due to obstacles present along the channel i.e. non-line of sight (NLOS) channel, rain, wind, and fog. In order to overcome these problems, several techniques have been proposed in [5-11], [13-15] but the obstruction of the transmitted laser beam due to the NLOS communication channel was not solved. For secure data transmission over a known turbulent environment, closed-form expressions for the secrecy outage probability, secrecy capacity, and the strictly positive secrecy capacity were derived and the data was received securely at the receiver end by preventing the eavesdropper's attack [5]. An orthogonal frequency division multiplexing (OFDM) based FSO wireless communication system that is more resistant to the rainy weather conditions and with a data rate of up to 10Gbps was proposed in [6].

A novel FSO communication system using Acousto-optic Modulator (AOM) as an external modulator to enhance the system performance is presented in [7]. To enhance the performance of the FSO wireless communication system under rainy weather conditions authors in [8] used the MIMO

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:56:08 UTC from IEEE Xplore. Restrictions apply.

technique and analyzed the signal-to-noise ratio (SNR) and Q-factor at the receiver end. The BER performance of MIMO and single input single output (SISO) under the atmospheric turbulence effect in the FSO wireless communication is presented in [9]. OFDM-FSO, WDM-FSO are the two widely used techniques that enhance the system performance with reference to high speed and long distances [10]. In this study, Build an 82-m 9 Gb/s four-level pulse amplitude modulation practically (PAM4)-FSO-plastic optical fiber (POF)-underwater wireless optical communication (UWOC) convergent system that successfully builds a long-haul wireless-wired optical link using, optical beam reducer, graded-index (GI)-plastic optical fiber (POF), and doublet lenses [11].

To overcome the aforementioned challenges related to NLOS transmission and atmospheric turbulence due to haze and rain, we propose a multiple-input single-output (MISO) FSO wireless communication system using multiple LASER beams for transmission of a single signal. In this paper, we introduce an "FSO Wireless Communication system using Array of LASERs for Data Transmission." Firstly, we design the MISO-FSO communication system model in OptiSystem 7 and Proteus 7 simulation software. Then, we present a real-time hardware setup for the same. We have used a converging lens set up just before the receiver input in the hardware system to converge all the laser beams, carrying the data, to the receiver input. In Optisystem, this particular step is implemented as a combination of power combiner and the lens. The proposed model is easy to implement, reliable, efficient, and of low cost. We also analyze the proposed system model in terms of eye-diagram, received power, BER, and Q-factor, with respect to (w.r.t.) link distance, under haze, and rainy conditions.

The paper is organized as follows; in section II, we discuss the MISO-FSO wireless communication system implementation. In section III, we discuss the attenuation due to different atmospheric conditions such as heavy rain, medium rain, light rain, heavy haze, and light haze. Results analysis and Conclusion is discuss in section IV and V respectively.

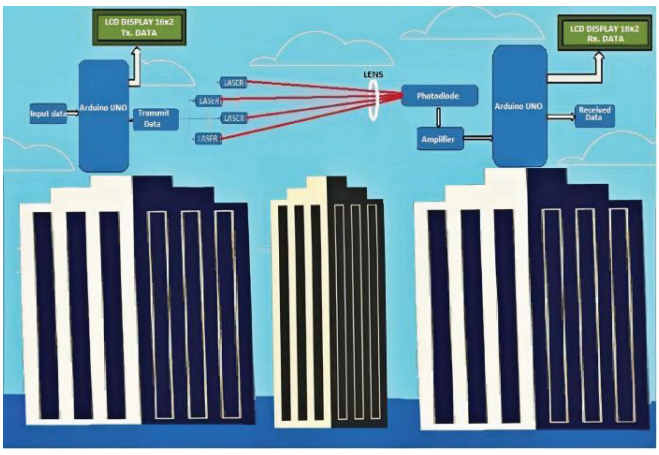


Fig. 1. Block diagram of MISO-FSO wireless communication system

II. FSO Wireless Communication System Implementation

In the first stage, we implement the prototype for the proposed MISO-FSO communication using the OptiSystem7 software. At the transmitter side, we use a pseudo sequence generator, LASER source, and non-return to zero (NRZ) modulator. We use photodiode, converging lens, amplifier, and 3R regenerator at the receiver end. The photodiode characteristics plot in Fig. 2 shows the relationship between the photocurrent and illumination intensity. The expression in (1) is used to determine the important parameter K i.e. photodetector sensitivity. The photosensitivity (K) for photodetector is defined as "the photodetector capability to differentiate an incident optical signal from the noise," [12].

$$K = \frac{I_{photo}}{I_{dark}} \tag{1}$$

where, I_{photo}=I_{ill}-I_{dark}, I_{photo} denotes photocurrent, I_{dark} denotes dark current, and I_{ill} indicates the illumination intensity. The photocurrent vs illumination intensity plot is shown in Fig. 2 [12].

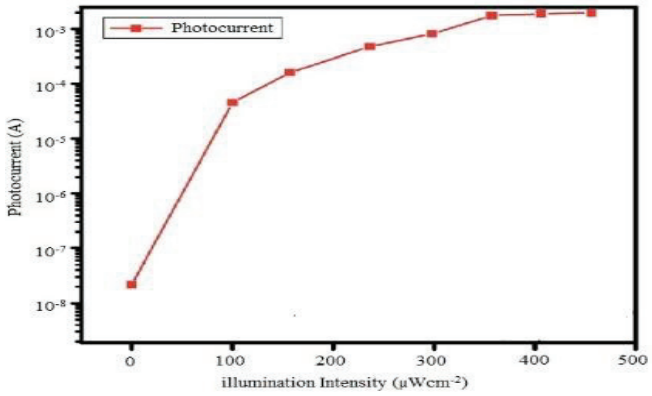


Fig. 2. Relationship between photocurrent vs illumination intensity [12]

By using the converging lens at the receiver side, the intensity is increased, and from Fig. 2. it can be seen that this will result in better detection at the receiver. We analyze the received power, BER, Q-factor and link distance with the eye diagram analyzer and optical-power-meter_10 as shown in Fig. 3. We have also simulated the transmitter and receiver section in Proteus software, as shown in Fig. 4. In this simulation, the data is transmitted from the transmitter side to the receiver side through free space to investigate that whether the proposed FSO scheme is feasible or not.

In the second stage, we implement the prototype of the real-time hardware setup and check the performance and reliability of the data communication. Because in FSO communication, the main problem is when an obstacle comes in contact with the line of sight (LOS) signal and interrupts its reception. So, here we use arrays of LASERs for uninterrupted data transmission by utilizing the spatial diversity at the transmitter side.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:56:08 UTC from IEEE Xplore. Restrictions apply.

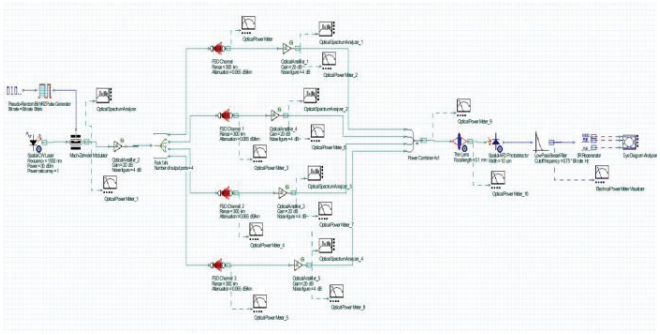


Fig. 3. Simulation setup for transmitter and receiver in OptiSystem7

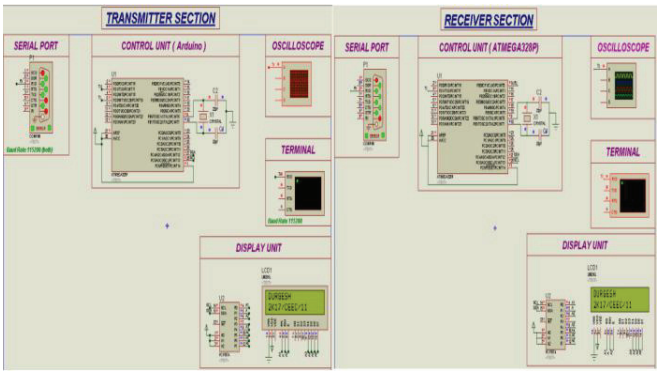


Fig. 4. Simulation setup for transmitter and receiver in Proteus 7

A. Simulation Setup in OptiSystem

We have used the OptiSystem 7 software to implement the simulation setup for the proposed FSO wireless communication system. For the simulation, we used a pseudo sequence generator, LASER source and NRZ modulator at the transmitter side and photodiode, power combiner, converging lens, amplifier, and 3R regenerator at the receiver end. Then we observed the eye diagram and analyzed it for light haze (see Fig. 5), heavy haze (see Fig. 6), light rain (see Fig. 7), medium rain (see Fig. 8), and heavy rain (see Fig. 9) weather conditions. The eye diagram shows the acceptable eye-opening to detect signals received on the receiver side. Then we analyze BER, Q-factor with the help of Eye diagram in various weather conditions.

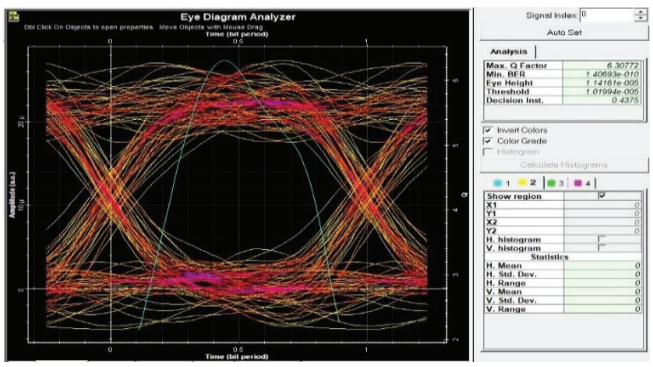


Fig. 5. Eye diagram for light haze at a distance 73.76 km

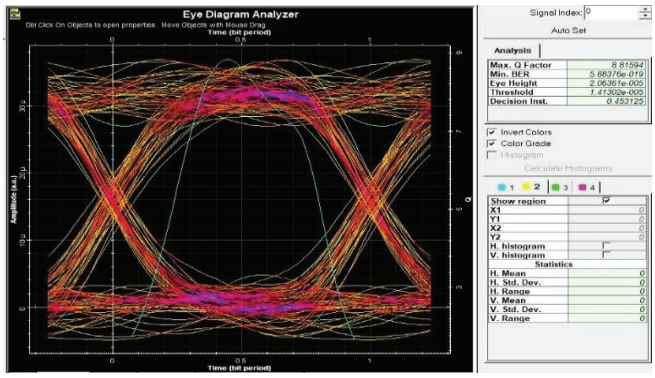


Fig. 6. Eye diagram for heavy haze at a distance 21.6 km

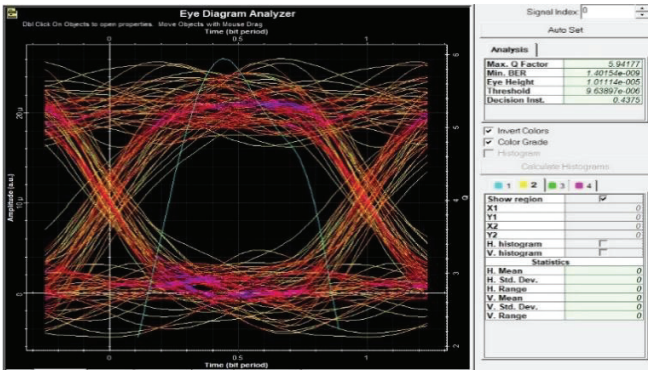


Fig. 7. Eye diagram for light rain at a distance 9.33 km

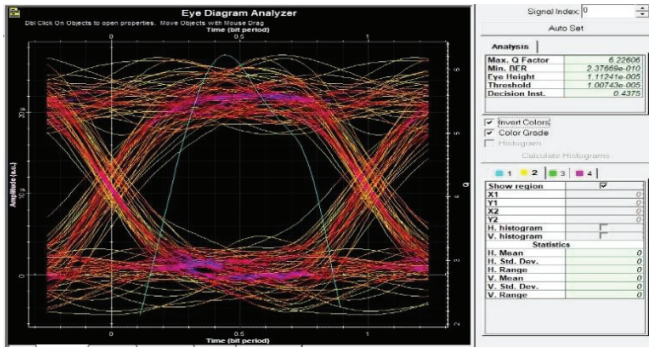


Fig. 8. Eye diagram for medium rain at a distance 6.38 km

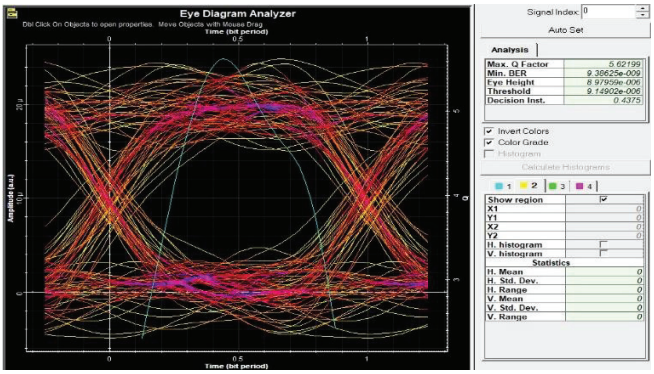


Fig. 9. Eye diagram for heavy rain at a distance 3.44 km

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:56:08 UTC from IEEE Xplore. Restrictions apply.

B. Simulation Setup in Proteus

We have used the Proteus 7 software to implement the simulation setup for the proposed wireless FSO communication system. In this simulation, the information is transmitted from the transmitter side to the receiver side through free space to investigate that whether the proposed FSO scheme is feasible or not. The transmitter and receiver sections are presented in Fig. 4. We used components like Atmega328P microcontroller, an array of LASERs with a wavelength of 1550 nm, power supply, 16x2 LCD display, and photodiode. During the transmission, serial monitor is used as an input data sender, and the input data is sent to the transmitter end using a computer or laptop. The information was successfully received at the receiver's side during the simulation and observed the waveform for transmitted data and received data are shown in Fig. 10 and Fig. 11.

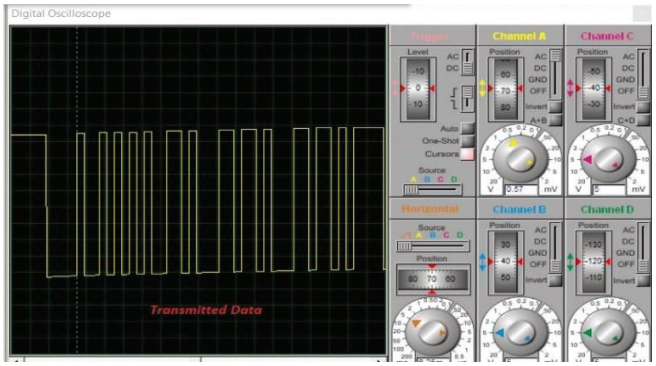


Fig. 10. Data transmitted at the transmitter end

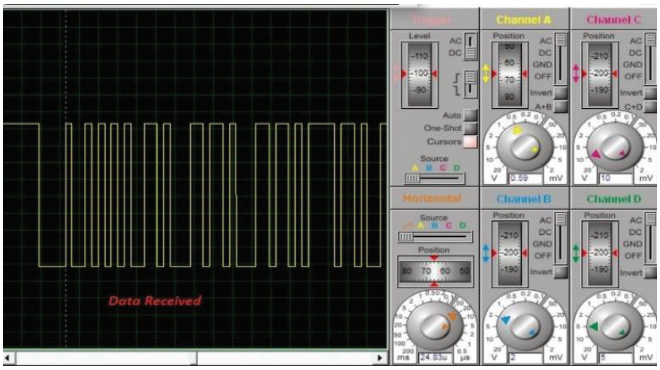


Fig. 11. Data received at the receiver end

C. Hardware Setup

In this section, we present a real-time hardware setup for the proposed MISO-FSO wireless communication system. The data transmitted from the transmitter side is presented in Fig. 12 and the data received by the receiver side is presented in Fig. 13. For designing the transmitter, we used components like Atmega328P microcontroller, power supply, array of LASERs with the wavelength of 1550 nm, power supply, 16x2 LCD display, photodiode, converging lens, and the serial monitor as an input data sender. The input data is sent at a baud rate of 115200 bits per second to the transmitter end by the serial monitor.

Here, we used four LASER sources and set them on different positions and angles to implement the MISO scheme; these LASER beams converge at the receiver end into a single beam by the converging lens, and then we tested the real-time performance up to a distance of 22.83 feet. It is important to note that the same information is transmitted through each LASER source. If any obstacle obstructs any one of the LASER beam still the information can be received at the receiver end due to some other unobstructed LASER beam transmitted from the array of LASERs. There are 4 lasers named as A, B, C, and D. The successful transmission of information with the proposed setup can be further explained as: The same message is transmitted through each of the four beams placed at distinct positions at the transmitters end. Suppose, if a bird comes in contact with beam C, the information can still be gathered from the laser beams A, B, and D. By utilizing this spatial diversity, the probability of data interruption is very low. So, this system is more reliable and low cost for free space wireless communication.

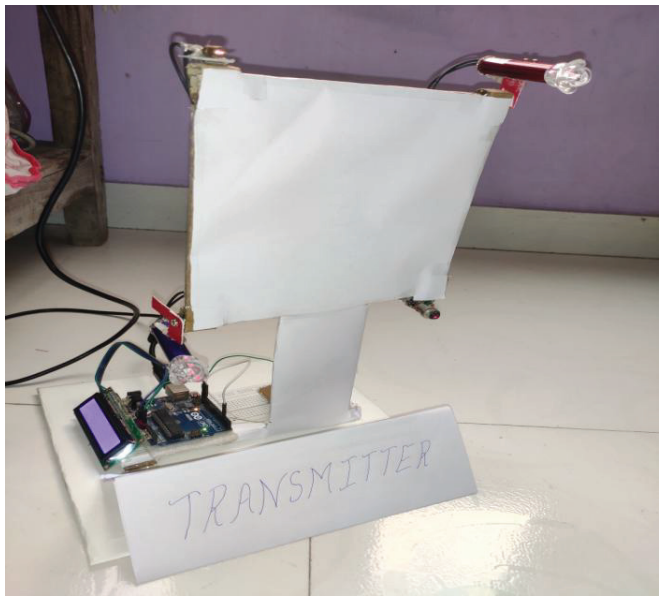


Fig. 12. Hardware setup for the transmitter



Fig. 13. Hardware setup for the receiver

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:56:08 UTC from IEEE Xplore. Restrictions apply.

III. Attenuation Due to Different Atmospheric Conditions

A. Attenuation in Rainy Season

The scattering of a LASER beam in different weather conditions is also one of the major problems which arise during FSO communication, and scattering results in attenuation of an optical signal. The attenuation factor depends on the amount or intensity of rainfall. As the rain increases, the attenuation factor increases. The optical association with the rain rate of *R* mm/hr produces specific constriction esteems [13]:

$$\alpha_{rain} = (1.076)\, R^{0.67} \text{ dB/km} \quad (2)$$

Here, *R* denotes rainfall rate, and α_{rain} denotes attenuation due to the rain.

Beer's law describes the laser attenuation for different weather conditions [14] as

$$\tau(R) = \left(\frac{Pr}{Pt}\right) = e^{-\beta R} \quad (3)$$

Here, $\tau(R)$ denotes transmittance at distance *R* (km), *R* denotes link distance, P_r denotes LASER power at distance *R*, P_t denotes power (Watt), β denotes scattering coefficient (km^{-1}).

Non-selective scattering exists when the particles' diameter in the atmosphere are larger than the radiation wavelength. It is due to droplets of water in the atmosphere. The scattering coefficient is determined using Stroke's theorem [15] as:

$$\beta = \pi\alpha^2 N_\alpha\, Q_{scat} \left(\frac{\alpha}{\lambda}\right) \quad (4)$$

Here, α denotes raindrop radius, $N\alpha$ denotes raindrop distribution, Q_{scat} denotes scattering efficiency, and λ denotes the wavelength. N_α is determined by the equation [15]:

$$N_\alpha = \frac{3Z\alpha}{4/3\pi\alpha^3 V\alpha} \quad (5)$$

Here, $Z\alpha$ denotes rainfall rate (cm/s), $V\alpha$ denotes maximum speed limit precipitation. The raindrop speed is calculated by using the equation [15]:

$$V_\alpha = \frac{2\alpha^2 \rho g}{9\eta} \quad (6)$$

Here, ρ denotes water density, η denotes air stickiness, and g denotes gravitational constant. The values for the parameters are presented in Table I, and the attenuation and link distance under various weather conditions are presented in Table II.

TABLE I. Value of Parameters [15]

Constant Value of Parameters		
S. No	***Constant***	***Value***
1	Water density , ρ	1 g/cm^2
2	gravitational constant, g	980 cm/s^2
3	Air stickiness, η	1.8 x 10^{-4} g/cm
4	Wavelength, λ	1550 nm
5	Q_{scat}	2
6	Rain droplet, α	0.0001-0.1 cm

B. Attenuation in Haze

The attenuation at various levels of beam visibility determines the FSO communication system performance. Because particles in the atmosphere stay longer in comparison to the rain due to haze. Therefore, creating deterioration in the performance of the FSO system. Onsite testing is conducted to assess the FSO communication system performance. If the system performance in terms of attenuation is satisfactory, then permanent installation of the system is done. Otherwise, if the results are not up to the mark, the parameters require changes and adjustment. The system performance can be analyzed by using Kim & Kruse Model [15] and given by haze attenuation as:

$$\beta = \frac{3.91}{V}\left(\frac{\lambda}{1550\text{ nm}}\right)^{-q} \quad (7)$$

Here, β denotes the haze attenuation, λ denotes wavelength, V denotes visibility (km), q denotes distribution size of scattering particles. Attenuation and link distance under various weather conditions without converging lens are presented in Table II and with the converging lens are presented in Table III.

TABLE II. Link Distance Under Various Weather Conditions without Converging Lens

Weather conditions	***Attenuation (dB/km)***	***Received Power (dBm)***	***Link distance (km)***
Light haze / fog	0.55	-13.273	48.50
Heavy haze	2.37	-15.095	16.00
Light rain	6.27	-16.9645	7.40
Medium rain	9.64	-13.6925	4.85
Heavy rain	19.28	-18.2841	2.90

TABLE III. Link Distance Under Various Weather Conditions with Converging Lens

Weather conditions	***Attenuation (dB/km)***	***Received Power (dBm)***	***Link distance (km)***
Light haze / fog	0.55	-24.979	73.76
Heavy haze	2.37	-24.949	21.60
Light rain	6.27	-24.978	9.33
Medium rain	9.64	-24,765	6.38
Heavy rain	19.28	-24.249	3.44

IV. Results Analysis

Detailed analysis of link distance in MISO-FSO wireless communication is done under different weather conditions. The rainy and hazy conditions are considered during heavy rain, medium rain, light rain, heavy haze, and light haze. Here, we observe the received power, Q-factor, and BER w.r.t. link distance between the transmitter and receiver.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:56:08 UTC from IEEE Xplore. Restrictions apply.

A. Rain

a) Received Power: The received power plot is shown in Fig. 14, representing the relationship between received power and link distance under heavy rain, medium rain, and light rain weather conditions. The power decreases with increasing distance. For the link distance of 9.33 km at the attenuation 6.27 dB/km, the power received is -24.978 dBm during the light rain weather condition. For the link distance of 6.38 km at the attenuation 9.64 dB/km, power received is -24.765 dBm during the medium rain weather condition. For the link distance of 3.44 km at the attenuation of 19.28 dB/km, the power received is -24.249 dBm during the heavy rain weather condition.

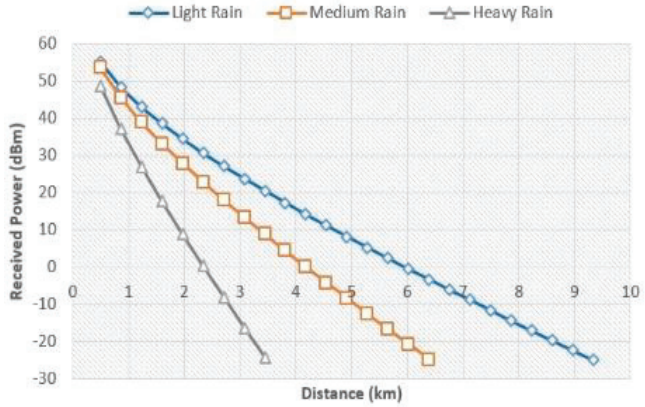


Fig. 14. Relationship between received power vs link distance

b) Quality factor (Q-factor): The Q-factor plot is shown in Fig. 15, representing the relationship between quality-factor and distance under heavy rain, medium rain, and light rain weather conditions. The Q-factor varies inversely w.r.t. the link distance. For the link distance of 9.33 km at the attenuation 6.27 dB/km, the Q-factor is 4.657 during the light rain weather condition. For the link distance of 6.38 km at the attenuation 9.64 dB/km, the Q-factor is 4.885 during the medium rain weather condition. For the link distance of 3.44 km at the attenuation 19.28 dB/km, the Q-factor is 5.48 during the heavy rain weather condition.

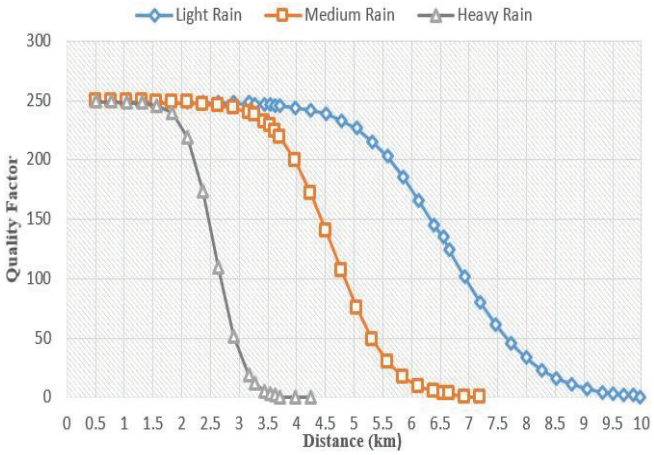


Fig. 15. Relationship between quality factor vs link distance

c) Bit Error Rate (BER): The BER plot is shown in Fig. 16, representing the relationship between BER and link distance under heavy rain, medium rain, and light rain weather conditions. The BER varies with respect to the link distance, which is quite obvious to understand. For the link distance of 9.33 km at the attenuation 6.27 dB/km, the BER is 1.15E-06 bits during the light rain weather condition. For the link distance of 6.38 km at the attenuation 9.64 dB/km, the BER is 1.16E-06 bits during the medium rain weather condition. For the link distance of 3.44 km at the attenuation 19.28 dB/km, the BER is 1.12E-05 bits during the heavy rain weather condition.

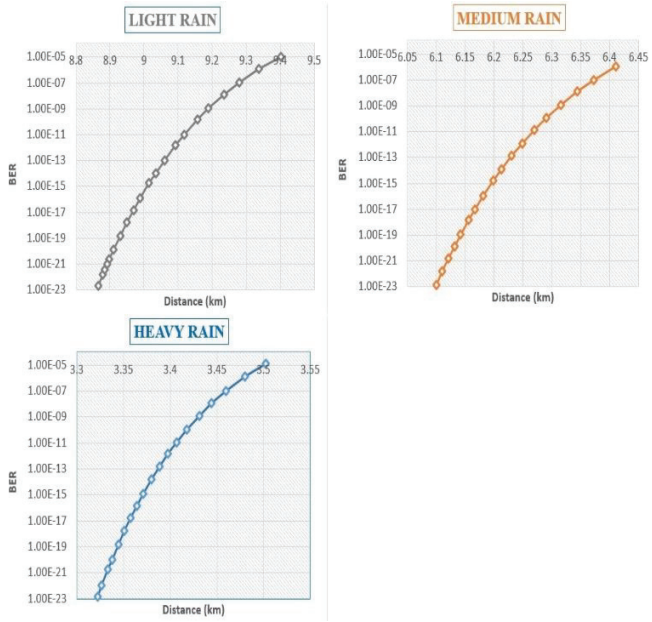


Fig. 16. Relationship between BER vs link distance

B. Haze

a) Received Power: During the light haze weather condition, for the link distance of 73.76 km at the attenuation 0.55 dB/km, the power received is -24.979 dB. During the heavy haze weather condition, for the link distance of 21.6 at the attenuation 2.37 dB/km, the power received is -24.949 dB.

b) Quality factor: During the light haze weather condition, For the link distance of 73.76 km at the attenuation 0.55 dB/km, the Q-factor is 4.77. During heavy haze weather condition, For the link distance of 21.6 km at the attenuation 2.37 dB/km, the Q-factor is 4.81.

c) Bit Error Rate: During the light haze weather condition, For the link distance of 73.76 km at the attenuation 0.55 dB/km, the BER is 9.02E-07 bits. During the heavy haze weather condition, For the link distance of 21.6 km link distance at the attenuation 2.37 dB/km, the BER is 7.43E-07 bits.

C. Real-Time Hardware Testing

We tested the performance of the hardware prototype for the proposed scheme up to a distance of 22.83 feet. For testing purposes, we transmitted a text paragraph through the serial

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:56:08 UTC from IEEE Xplore. Restrictions apply.

terminal with a baud rate of 115200 bits per second as shown in Fig. 17. Information was sent through the transmitter end and successfully received without any error at the receiver's end at the serial terminal with the help of a photodiode as shown in Fig. 18. This implementation verifies the reliability of the proposed MISO-FSO wireless communication system under a realistic environment. This hardware setup can further be scaled up to a bigger size system to implement the proposed FSO communication system with high data rate and reliability for a larger link distance.

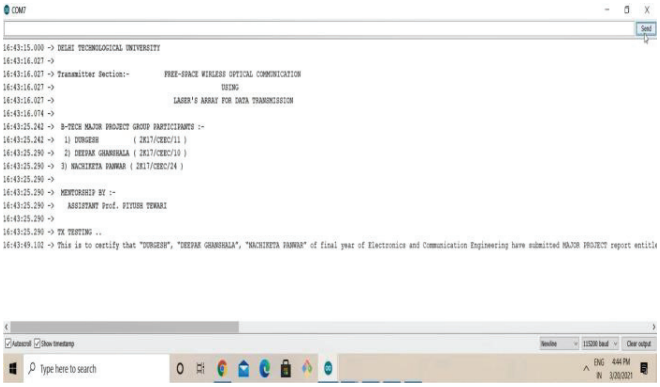


Fig. 17. Transmitted data by serial monitor

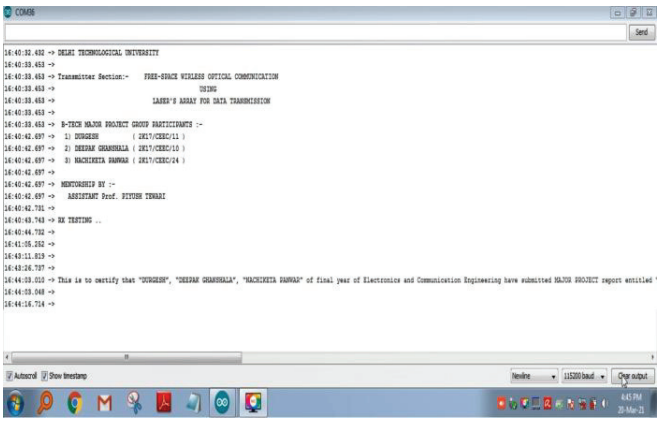


Fig. 18. Received data by serial monitor

V. Conclusion

In this paper, we have used Li-Fi for a high-speed MISO-FSO wireless communication system using an array of LASERs for the transmission of the same data signal. We analyzed the proposed system performance in terms of Q-factor, BER, received power, eye diagram with respect to the link distance. For this study, 1550 nm LASER wavelength and 20 Gb/s multi-beam transmission were used to achieve secure transmission successfully. Reliable communication performance up to link distances of 3.44 km, 6.38 km, 9.33 km, 21.6 km, and 73.76 km were successfully achieved under heavy rain, medium rain, light rain, heavy haze, and light haze weather conditions. On the basis of obtaining results, we can say that during the heavy rain condition in areas such as mawsynram, cherrapunji and pondicherray we can get reliable performance achieved using the proposed MISO-FSO scheme up to link distance 3.44 km.

A hardware prototype for the proposed scheme was also developed, and the communication performance was analyzed. The array of LASERs and the converging lens work as an OR gate as when all the transmitted LASER beams are interrupted, then no information is received at the receiver end, otherwise, information is successfully received at the receiver end. So, this method decreases the probability of data interruption due to the NLOS channel condition. The converging lens also increases the LASERs beam's intensity at the receiver side and focuses all the beams on the photodiode at the receiver.

References

[1] Harald Haas and Tezcan Cogalan, "LiFi opportunities and Challenges," 16th International Symposium on Wireless Communication Systems, IEEE, 2019.

[2] AL KLASCIUS, "Microwave radiation hazards around large microwave antenna," American Industrial Hygiene Association Journal, 2012.

[3] Xu-Hong Huang, Chung-Yi Li, Hai-Han Lu, Chung-Wei Su, You-Ruei Wu, Zhen-Han Wang and Yong-Nian Chen, "WDM free-space optical communication system of high speed hybrid signals," IEEE, 2018.

[4] Toshimasa Umezawa, Takahide Sakamoto, Atsushi Kanno, Atsushi Matsumoto, Kouichi Akahane, NaoKatasu Yamamoto, and Tetsuya Kawanishi, "25-Gbaud 4-WDM free-space optical communication using high-speed 2D photodetector array," IEEE, 2018.

[5] Mohammad Javed Saber and Seyed Mohammad Sajad Sadough, "On secure free space optical communications over malaga turbulence channels," IEEE, 2017.

[6] Fakhriy Hario, Eka Maulana, Sholeh Hadi Pramono, Sapriest Nainy and Anas Machfudy Al Junaedi, "Design of OFDM-FSO commuication system on high data rate for tropical climate region," 2019 International Conference on Advanced Technologies for Communications (ATC), IEEE, 2019.

[7] Raed Mesleh, and Ayat AL-Olaimat, "Acousto-optical modulators for free space optical wireless communication systems," IEEE, 2018.

[8] Farouk Kh. Shaker and Mazin Ali A. Ali, "Multi-beam free-space optical link to mitigation of rain attenuation," Journal of Optical Communications, EISSN 2191-6322, ISSN 0173-4911, Walter de Gruyter GmbH, 2018.

[9] K. Prabu, "Analysis of FSO systems with SISO and MIMO techniques," Springer Professional, 2019.

[10] Aditi Malik and Preeti Singh, "Free space optics: current application and future challenges," Hindawi, 2015.

[11] Chung-Yi Li, Hai-Han Lu, IEEE Senior Member, Yun-Chieh Wang Zhen-Han Wang, Chung-Wei Su, Yi-Feng Lu and Wen-Shing Tsai, "An 82-m 9 Gb/s PAM4 FSO-POF-UWOC convergent system," Vol 11, IEEE, 2019.

[12] Muhammad Sulaman, Shengyi Yang, Taojian Song, Haowei Wang, Yishan Wang, Bo He, Miao Dong, Yi Tang, Yong Song and Bingsuo Zou, "High performance solution processed infrared photodiode based on ternary PbsSxSe1-x colloidal quantum dots." RSC Advances, RSC Adv., 2016, 6, 87730.

[13] Farukh Nadeem, Vaclav Kvicera, Muhammad Saleem Awan, Erich Leitgeb, Sajid Sheikh Muhammad and Gorazd Kandus, "Weather effects on hybrid FSO/RF communication link," IEEE Journal on Selected Areas in Communication, Vol. 27, No.9, IEEE, 2009.

[14] B.S. Naimullah, M. Othman, A. K. Rahman, S.I. Sulaiman, S. Ishak, S. Hitam and S.A. Aljunid, "Comparison of wavelength propagation for free space optical communications," 2008 International Conference on Electronic Design, IEEE, 2008.

[15] Hilal A. Fadhil, Angela Amphawan, Hasrul A.B. Shamsuddin, Thanaa Hussein Abd, Hamza M.R. Al-Khafaji, S.A. Aljunid and Nasim Ahmed, "Optimization of free space optics parameters: an optimum solution for bad weather conditions," JLEO-52845, Int. J. Light Electron Opt., 2012.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:56:08 UTC from IEEE Xplore. Restrictions apply.

arXiv:2108.10639v1 [stat.ML] 24 Aug 2021

GRADE: A GRAPH BASED DATA-DRIVEN SOLVER FOR TIME-DEPENDENT NONLINEAR PARTIAL DIFFERENTIAL EQUATIONS

A PREPRINT

Yash Kumar
Department of Mechanical Engineering
Delhi Technological University
yashk8481@gmail.com

Souvik Chakraborty
Department of Applied Mechanics
School of Artificial Intelligence (ScAI)
India Institute of Technology (IIT) Delhi
souvik@am.iitd.ac.in

August 25, 2021

ABSTRACT

The physical world is governed by the laws of physics, often represented in form of nonlinear partial differential equations (PDEs). Unfortunately, solution of PDEs is non-trivial and often involves significant computational time. With recent developments in the field of artificial intelligence and machine learning, solution of PDEs using neural network has emerged as a domain with huge potential. However, most of the developments in this field are based on either fully connected neural networks (FNN) or convolutional neural networks (CNN). While FNN is computationally inefficient as the number of network parameters can be potentially huge, CNN necessitates regular grid and simpler domain. In this work, we propose a novel framework referred to as the Graph Attention Differential Equation (GrADE) for solving time dependent nonlinear PDEs. The proposed approach couples FNN, graph neural network, and recently developed Neural ODE framework. The primary idea is to use graph neural network for modeling the spatial domain, and Neural ODE for modeling the temporal domain. The attention mechanism identifies important inputs/features and assign more weightage to the same; this enhances the performance of the proposed framework. Neural ODE, on the other hand, results in constant memory cost and allows trading of numerical precision for speed. We also propose depth refinement as an effective technique for training the proposed architecture is lesser time with better accuracy. The effectiveness of the proposed framework is illustrated using 1D and 2D Burgers' equation. Results obtained illustrate the capability of the proposed framework in modeling PDE and its scalability to larger domains without the need for retraining.

***K*eywords** Graph Neural Network · Attention · Neural ODE · PDE · non-linearity

1 Introduction

Many complex phenomena of scientific importance can be compressed into a few partial differential equations (PDEs). Solving them is key to understand these phenomena. Popular methods for solving PDEs include Finite Element Method [1], Finite Volume Method [2], Finite Difference Method [3], and Boundary Element Method [4]. However, these methods are often computationally expensive and can take hours, if not days, to solve complex nonlinear PDEs on irregular domains. Therefore, even today, development of efficient methods for solving PDEs is a relevant problem.

With recent developments in the field of artificial intelligence and machine learning, data driven solution of PDEs has emerged as a possible alternative to the classical numerical techniques. The primary idea of these methods is to learn the dynamical evolution by using machine learning algorithms. Popular machine learning algorithms used for learning system dynamics include reduced-order models [5, 6], polynomial chaos expansion [7], and Gaussian processes [8, 9]. Others have tried to find governing equation using symbolic regression from data [10, 11]. Brunton et al. [12] and Raissi

and Karniadakis [13] also attempted to obtain equations that best describes the observed data. Patel and Desjardins [14] used neural networks over fourier transforms regressing nonlinear operator in PDE.

In past decade, research has been focused around using FNN, RNN for incrementing dynamics of system. Popular approaches includes models based on Long Short Term Memories and transformers [15, 16, 17]. Geneva and Zabaras [18] used physics constrained auto-regressive model for surrogate modeling of dynamical systems. Although these models have been effective at modeling the systems dynamics, there is a lack of transparency as they act as black box models. Moreover these are discrete models and predicts sequence separated with only fixed time step. On the other hand, models used in [19, 12] works with equation, but requires numerical time derivative of data; this naturally becomes a potential source of error.

Another popular class of methods for modeling nonlinear dynamical systems is the physics-informed neural network (PINN). The idea was initially applied to simple fully connected [20] and later extended to deep neural nets [21]. The basic idea here is to place a neural network prior on the state variable and then estimate the neural network parameters by using a physics-informed loss function. Continuous time and discrete time formulations of PINN were proposed. PINN has been successfully applied to solve a wide array of dynamical systems including, but not limited to, fluid flow [22], heat transfer [23], fracture mechanics [24], reliability analysis [25] and bio-mechanics [26]. Several improvements to the originally proposed PINN can also be found in the literature. For example, Zhu et al. [27] developed convolutional PINN for time-independent systems. Geneva and Zabaras [18] developed an auto-regressive convolutional PINN for dynamical systems. A Bayesian variant of the same was also proposed. In both the works, discrete time variants of PINN were used. The primary advantage of PINN resides in the fact that no training data is needed. However, the physics-informed loss function involved in PINN is difficult to optimize. Also, PINN assumes that the governing PDEs are exact, which is often not true. A few research directed towards addressing this issue can be found in the literature [28, 29].

Success of ResNet [30] in computer vision has attracted attention of researcher as it resembles Euler's time integration scheme and thus, introduces a bias in network architecture. Recently developed Neural ODE [31] extend this idea to more advanced integration schemes. These networks parameterize a differential equation as

$$u_t = \boldsymbol{f}(x,t,u(x,t);\boldsymbol{\theta}) \tag{1}$$

where, dynamic function $\boldsymbol{f} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^n$ and initial value $y_0 \in \mathbb{R}^n$. A NODE having one hidden unit encounters a problem where activation trajectories do not cross with depth of network limiting expressiblity of network. This is overcome by adding auxiliary dimensions [32]. There are various methods used for training NODE depending upon requirements. Besides usual auto-differentiation, adjoint-based back-propagation is used due to being memory efficient, but requires more time steps. Some challenges are overcome by using checkpoint method [33, 34].

In this work, we propose a novel framework, referred to as Graph Attention Differential Equation (GrADE) for learning system dynamics from data. The primary motivation behind GrADE resides in the fact that real-life data are unstructured and resides on irregular domains. This prohibits direct application of convolutional neural network based approaches. GrADE combines Graph Neural Network (GNN) with Neural ODE. With GNN, one can easily handle unstructured data on a irregular domain. Within GrADE, GNN is used to model the spatial domain and Neural ODE is used to model the temporal domain. Among different GNN avaialble in the literature, we propose to use the graph attention (GAT) [35] within the proposed GrADE. Within GAT, attention mechanism is used on embedding of nodes during aggregation. Additionally, GrADE allows a streamlined way of embedding the boundary conditions of required solutions in the architecture of the graph connections on boundary nodes. This ensures the network prediction always meet the required boundary conditions. Example for the same are discussed in the paper.

The rest of the paper is organized as follows. In Section 2, the problem statement has been defined. Brief review of fully connected neural network (FNN), attention mechanism, and GAT are presented in Section 3. We present the proposed approach in Section 4. Section 5 presents two examples to illustrate the applicability of the proposed approach. Finally, Section 6 presents the concluding remarks.

2 Problem statement

In this work, we are interested in discovering PDE using data. Without loss of generality, we consider a system governed by the following system of PDEs

$$\begin{aligned} \boldsymbol{u}(\boldsymbol{x},t)_t = \boldsymbol{f}(\boldsymbol{x},\boldsymbol{u}(\boldsymbol{x},t)),\ \boldsymbol{x} \in \Omega,\ t \in [0,T] \\ \boldsymbol{B}(\boldsymbol{u}) = \boldsymbol{b}(\boldsymbol{x},t),\ \boldsymbol{x} \in \Gamma \end{aligned} \tag{2}$$

where $\boldsymbol{u}(\boldsymbol{x},t) \in \mathbb{R}^{ndim}$ are the state variables, $\boldsymbol{u}(\boldsymbol{x},t)_t$ is temporal derivative and $\boldsymbol{B}$ is operator for enforcing boundary conditions. $\boldsymbol{x} \in \Omega$ represents the spatial coordinates and $t \in [0,T]$ represents time. Initial state $\boldsymbol{u}(\boldsymbol{x},0)$ can be any real valued random field.

We assume that we have noisy measurements of the state variables at fixed time intervals. Data is of form $\mathscr{D} = \{\boldsymbol{U}(\boldsymbol{x},t_i)\}_{i=1}^{N_t}$ where, N_t is the number of time-steps at which data is available. $\boldsymbol{U}(\boldsymbol{x},t_i)$ is vector representing the measurements of state at predefined fixed grid nodes at time t_i. We are interested in developing a framework that is able to predict the future evolution of the state variables $\boldsymbol{u}(\boldsymbol{x},t)$ and $\boldsymbol{u}(\boldsymbol{x},t)_t$. In other words, we are interested in learning the operator $\boldsymbol{f}(\cdot)$ in Eq. (2).

Remark 1: Time evolution of the state variable $\boldsymbol{u}(\boldsymbol{x},t)$ can be easily learned by using ResNet [36] or other similar framework. However, predicting time-evolution of $\boldsymbol{u}(\boldsymbol{x},t)_t$ is non-trivial as no measurements for the same is available. One can always opt for time-derivative. Finite difference type schemes results in erroneous results because of the noise in the data.

3 Brief review of feed-forward and graph neural network

In this section, we briefly review the fundamentals of Feed-forward Neural Network (FNN), attention mechanism and Graph Attention (GAT). These three form the backbone of the proposed approach.

3.1 Feed-forward neural network

One of the key component of the proposed GrADE is fully connected feed-forward neural network (FNN) also known as multilayer perceptron. FNNs are universal approximator [37] and are extremely accurate in performing a wide array of tasks such as statistical pattern recognition, regression and classification. Consider a $\mathscr{N}_N : \mathbb{R}^{N_{in}} \mapsto \mathbb{R}^{N_{out}}$ to be a operator of a FNN. Considering $\boldsymbol{x}_{in} \in \mathbb{R}^{N_{in}}$ to be the input, the output $\boldsymbol{x}_{out} \in \mathbb{R}^{N_{out}}$ can be represented as

$$\boldsymbol{x}_{out} = \mathscr{N}_N(\boldsymbol{x}_{in};\boldsymbol{\theta}), \tag{3}$$

where $\boldsymbol{\theta}$ represents the parameters of the neural network operator $\mathscr{N}_N$. In essence, the neural network operator $\mathscr{N}_N$ is composition function of the form

$$\mathscr{N}(\cdot;\boldsymbol{\theta}) = (\boldsymbol{\sigma}_M \circ \mathbf{W}_{M-1}) \circ \cdots \circ (\boldsymbol{\sigma}_2 \circ \mathbf{W}), \tag{4}$$

where $\mathbf{W}_j$ is the weight matrix connecting layer j and $(j+1)$, $\circ$ is operator composition, and $\boldsymbol{\sigma}_j : \mathbb{R} \mapsto \mathbb{R}$ is the activation function corresponding to the $j-$th layer. Note that the activation function is applied on one component at a time. The choice activation plays an important role in neural network. Popular activation functions available in the literature includes sigmoid, tan-hyberbolic, and rectified linear unit. Details on the activation function used in this paper is provided later.

For using a FNN in practice, one needs to estimate the parameters of $\mathscr{N}_N(\cdot;\boldsymbol{\theta})$. This is generally achieved by maximizing the likelihood of the data (or minimizing an error function). Considering $\mathscr{D}_d = \left\{\boldsymbol{x}_{in}^{(i)}, \boldsymbol{x}_{out}^{(i)}\right\}_{i=1}^{N_d}$ to be the training data available, we can estimate the parameters $\boldsymbol{\theta}$ by minimizing the $\mathscr{L}_2$ loss-function,

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \sum_{j=1}^{N_d} \left\| \boldsymbol{x}_{out}^{(j)} - \mathscr{N}_N(\boldsymbol{x}_{in}^{(j)};\boldsymbol{\theta}) \right\|^2, \tag{5}$$

where $\|\cdot\|$ represents the $\mathscr{L}_2$ norm. Note that other loss-functions like $\mathscr{L}_1$ norm can also be used.

Remark 2: FNN, although universal approximators, can potentially be computationally expensive. This is because all neurons at layer j are connected to all neurons at layer $j+1$. Therefore, the number of parameters in FNN is quite high.

3.2 Attention

Another core component of the proposed GrADE is the attention mechanism. In a conventional neural network, the hidden activation function $\sigma(\cdot)$ acts on a linear combination of the input activation. For instance, if $\boldsymbol{h}_i$ is the hidden state and $\boldsymbol{w}_i$ represents the weight, in a conventional neural network, we have

$$\boldsymbol{h}_{i+1} = \boldsymbol{\sigma}\left(\boldsymbol{w}_i^T \boldsymbol{h}_i\right), \tag{6}$$

where $\boldsymbol{h}_{i+1}$ is the output of the $i-$th layer. Note that the $\boldsymbol{w}_i$ in Eq. (6) is constant. In attention mechanism, we take a different path where the weight vectors are dependent on the inputs. This is mathematically represented as

$$\boldsymbol{h}_{i+1} = \boldsymbol{\sigma}(\boldsymbol{g}(\boldsymbol{h}_i;\boldsymbol{\alpha})^T \boldsymbol{h}_i), \tag{7}$$

where $\boldsymbol{g}(\cdot;\boldsymbol{\theta})$ represents a learnable function parameterized by parameters $\boldsymbol{\alpha}$. With such a setup, we are forcing the neural network to "pay attention" to different type of inputs in an adaptive manner.

The basic idea of attention was first proposed in the context of recurrent neural network (RNN); however the concept is equally applicable to other types of neural networks as well. For instance, [38, 39] proposed soft attention mechanism where the context vector in the decoder function of RNN is allowed to be function of input encoding vectors. On the other hand, [40] used attention mechanism within the convolutional neural network framework. Recently, PINN based on attention mechanism has also been developed [41]. Motivated from [35], we utilize attention mechanism within the graph neural network in this paper.

Researchers over the past few years have proposed different attention mechanism. For example, the attention mechanism shown in Eq. (7) is a type of multiplicative attention unit. The most popular attention mechanism are perhaps the 'dot product attention' and the 'multi-head attention'. Transformers [42], for instance, utilizes multi-head attention. Similarly, 'soft' and 'hard' attention mechanism can also be found in the literature. In hard attention, each output only attends one input location. However, with such a setup, the loss-function of neural network becomes non-differentiable. In this work, we use a soft attention mechanism. For further detauils on different attention mechanism, interested readers may refer [42, 43].

3.3 Graph neural network with attention

Having discussed the attention mechanism, we proceed to the last component of proposed GrADE, namely Graph Attention (GAT) [35]. We first briefly discuss graph neural network followed by GAT.

We define a graph $\mathscr{G} = \{\mathscr{V}, \mathscr{E}\}$ having vertices $\mathscr{V} = \{v_1, v_2, \ldots, v_N\}, N = |\mathscr{V}|$ and edges $\mathscr{E} \subseteq \mathscr{V} \times \mathscr{V}$. An edge in a graph connects two vertices and is denoted as $e_{i,j} := (v_i, v_j) \in \mathscr{V} \times \mathscr{V}$ with $1 \leq i, j \leq N$ and $i \neq j$. At this stage, we note that most real-world data lies on irregular domains (e.g., social networks, point cloud, biological networks) and can be represented using graphs. With graph neural network, it is possible to directly operate on these graphs. There are two major class of methods used for working with graphs. First is spectral methods which is based on spectral graph theory. It relies on convolution theorem for defining convolution on graph. Where Fourier transform of function on graph is performed via projecting the function on Fourier functions which are nothing but matrix of Eigenvectors of graph Laplacian obtained via expensive Eigen-decomposition. Moreover their is no guarantee of learning spatially localized filter. Henaff et al. [44] used linear combination of smooth kernels to approximate spectral filter resulting in localized spacial filters and smaller number of parameters. Defferrard et al. [45] and Levie et al. [46] used Chebyshev and Cayley's expansion for estimating spectral filter, bypassing the expensive Eigen-decomposition. Second is spatial methods which applies convolution directly on graph. Scarselli et al. [47] introduced the vanilla GCNs in which nodes shared same weights with all neighbours. This approach can handle different neighborhood sizes and is independent of graph size. Recently developed GraphSAGE [48] differentiate between weights of central node from neighbours while sampling from neighborhood, this feature improves performance of the model over various inductive benchmarks. Unlike classic convolution nets, this setting is isotropic in nature and do not distinguish between neighbours. Anisotropy can be achieved naturally if we have edge feature or using mechanism differentiating neighbours. MoNets [49] leverages Bayesian Gaussian mixture model parameters to differentiate between neighbours based on information about degree of node. GAT [35] used attention mechanism on node embeddings during aggregation.

Similar to convolutional neural networks, graph neural networks are composed of stacked layers, each performing message-passing and propagation. Most basic type of graph neural network layer can be represented vectorially as

$$\boldsymbol{h}_i^{l+1} = \sigma\left(\frac{1}{d_i}\sum_{j \in N_i} \boldsymbol{A}_{ij}\boldsymbol{W}^l\boldsymbol{h}_j^l\right), \tag{8}$$

where, $\boldsymbol{h}_i^{l+1} \in \mathbb{R}^d$ has a dimensions of $d \times 1$. Eq. (8) represents the operation performed on each node $v \in \mathscr{V}$ while implementing a layer of graph neural network. Note that the summation in Eq. (8) is carried out over the N_i neighbors of the $i-$th node.

Veličković et al. [35] used attention mechanism for message passing on graph-structured data. This approach attend to neighbors based upon attention weights and were able to achieve state-of-the-art results on Cora, Citeseer and Pubmed citation network datasets. In GAT, Eq. (8) is modified as follows

$$\boldsymbol{h}_i^{l+1} = \Big\Vert_{k=1}^{K} \left(\sigma\left(\sum_{j \in N_i} e_{ij}^{k,l}\boldsymbol{W}^{k,l}\boldsymbol{h}_j^l\right)\right), \tag{9}$$

where $\Vert$ represents concatenation, $\boldsymbol{W}^{k,l}$ is the weight matrix for input linear transformation, and $e_{ij}^{k,l}$ are the normalized attention coefficient and computed as

$$e_{ij} = \frac{\exp\left(\sigma'\left(\boldsymbol{\alpha}^T\left[\boldsymbol{W}\boldsymbol{h}_i^l \Vert \boldsymbol{W}\boldsymbol{h}_j^l\right]\right)\right)}{\sum_{k \in N_i}\exp\left(\sigma'\left(\boldsymbol{\alpha}^T\left[\boldsymbol{W}\boldsymbol{h}_i^l \Vert \boldsymbol{W}\boldsymbol{h}_k^l\right]\right)\right)}. \tag{10}$$

Note that σ' in Eq. (10) represents the activation functions and LeakyReLU is a popular choice in this case. $\boldsymbol{\alpha}$ in Eq. (10) represents the parameters of $\boldsymbol{f}(\cdot)$ defined in Eq. (7).

4 Proposed approach

In this section, we discuss the proposed framework referred to here as Graph Attention Differential Equation (GrADE) for learning dynamics of systems from data. Recall that given data $\mathscr{D} = \{\boldsymbol{x}, t_i, \boldsymbol{U}(\boldsymbol{x}, t_i)\}_{i=1}^{N_t}$ at fixed time interval, the objective here is to learn the operator $\boldsymbol{f}(\cdot)$ in Eq. (2); this will allow predicting the state variable $\boldsymbol{u}(\boldsymbol{x}, t)$ and its derivative $\boldsymbol{u}(\boldsymbol{x}, t)_t$ at future time-steps. We note that unlike other similar works existing in the literature [31, 50], the state variable $\boldsymbol{u}$ for our case is dependent on both spatial location and temporal location and hence, both spatial and temporal discretization will be required.

We proceed by placing a neural network prior to parameterize the differential equation in Eq. (2),

$$\boldsymbol{u}_t = \mathscr{N}_N(\boldsymbol{x}, t, \boldsymbol{u}(\boldsymbol{x}, t); \boldsymbol{\theta}) \tag{11}$$

where, $\boldsymbol{u}_t$ is time derivative of state vector $\boldsymbol{u}$, $\boldsymbol{\theta}$ are parameters of network $\mathscr{N}$. We rewrite Eq. (11) as

$$\boldsymbol{u}(\boldsymbol{x}, t_{k+1}) = \boldsymbol{u}(\boldsymbol{x}, t_k) + \int_{t_k}^{t_{k+1}} \mathscr{N}_N(\boldsymbol{x}, t, \boldsymbol{u}(\boldsymbol{x}, t)) dt. \tag{12}$$

Eq. (12) can be solved using some time integration scheme; although, time integration scheme introduces discretization error into the solution. For example, if we use Euler scheme, the approximation error is of the order $\mathscr{O}(\Delta t)$, where Δt is the time-step. In this work, we have used fourth order Runge-Kutta (RK4 - 3/8) scheme,

$$\boldsymbol{y}_1 = \mathscr{N}_N(\boldsymbol{x}, t_k, \boldsymbol{u}(\boldsymbol{x}, t_k)), \tag{13a}$$

$$\boldsymbol{y}_2 = \mathscr{N}_N\left(\boldsymbol{x}, t_k + \frac{\Delta t}{3}, \boldsymbol{u}(\boldsymbol{x}, t_k) + \left(\frac{\boldsymbol{y}_1}{3}\right)\right), \tag{13b}$$

$$\boldsymbol{y}_3 = \mathscr{N}_N\left(\boldsymbol{x}, t_k + \frac{2\Delta t}{3}, \boldsymbol{u}(\boldsymbol{x}, t_k) - \left(\frac{\boldsymbol{y}_1}{3} - \boldsymbol{y}_2\right)\right), \tag{13c}$$

$$\boldsymbol{y}_4 = \mathscr{N}_N(\boldsymbol{x}, t_k + \Delta t, \boldsymbol{u}(\boldsymbol{x}, t_k) + (\boldsymbol{y}_1 - \boldsymbol{y}_2 + \boldsymbol{y}_3)). \tag{13d}$$

$$\boldsymbol{u}(\boldsymbol{x}, t_{k+1}) = \boldsymbol{u}(\boldsymbol{x}, t_k) + \frac{\Delta t}{8}(\boldsymbol{y}_1 + 3\boldsymbol{y}_2 + 3\boldsymbol{y}_3 + \boldsymbol{y}_4) \tag{13e}$$

We assume the system of interest is autonomous, i.e., $\mathscr{N}_N$ does not explicitly depend on the temporal variable t. Therefore, network will only take previous state and spatial coordinate as input and need not vary with depth. With slight abuse to terminology, we here refer to steps involved in time-integration scheme as depth.

Remark 3: The idea of parameterizing the differential equation by a neural network is motivated from Neural ODE [31] and continuous-in-depth network [50]. However, both Neural ODE and continuous-in-depth network deals with ordinary differential equation. In our case the governing equation is a PDE.

To address the challenge mentioned in remark 3, we propose to use graph neural network to parameterize the operator $\boldsymbol{f}(\cdot)$ on the spatial domain. Accordingly, the neural network operator, $\mathscr{N}_N(\cdot)$ in Eq. (13) is to be replaced with $\mathscr{GN}(\cdot)$. The advantage of graph neural network resides in the fact that, unlike FNN, it only utilizes information from neighboring nodes; this makes the model computationally tractable and scalable. Additionally, graph neural network also generalizes well on unseen spatial domains. To be specific, we design a custom graph attention (GAT) network for approximating the operator in the spatial domain, Details on the custom GAT network proposed in this paper are discussed next.

4.1 GAT architecture

For approximating the operator $\boldsymbol{f}(\cdot)$ specified in Eq. (2) in the spatial domain, we consider a network consisting of two graph network layers. For building the custom graph network, we consider the followings:

- **Connection**: A node v_i in the graph is connected to its neighboring nodes. For 1D problem, we consider 4 nearest nodes to be neighbors. Similarly for 2D problem, we consider 8 neighboring nodes to be neighbors. This is schematically shown in Fig. 1. We use k-nearest neighbor algorithm [51] for determining neighbors of a node. Note that this will yield erroneous graph connection for the boundary nodes. In this work, the boundary nodes were modified manually.

- **Boundary conditions:** Boundary conditions (BC) are met by altering connection between graph nodes. e.g. for Dirichlet BC, we can remove the edges going towards the boundary nodes. This will prevent the value of boundary nodes from changing in next time step. We can specify new value for boundary at each time. Similarly for Neumann BC, we first remove edges going towards the boundary nodes and compute the boundary state variables based on the neighboring nodes by using the Taylors' series expansion. In this work, we consider examples with Periodic BC in which boundary nodes are connected both with local neighbours and nodes on opposite side of domain.

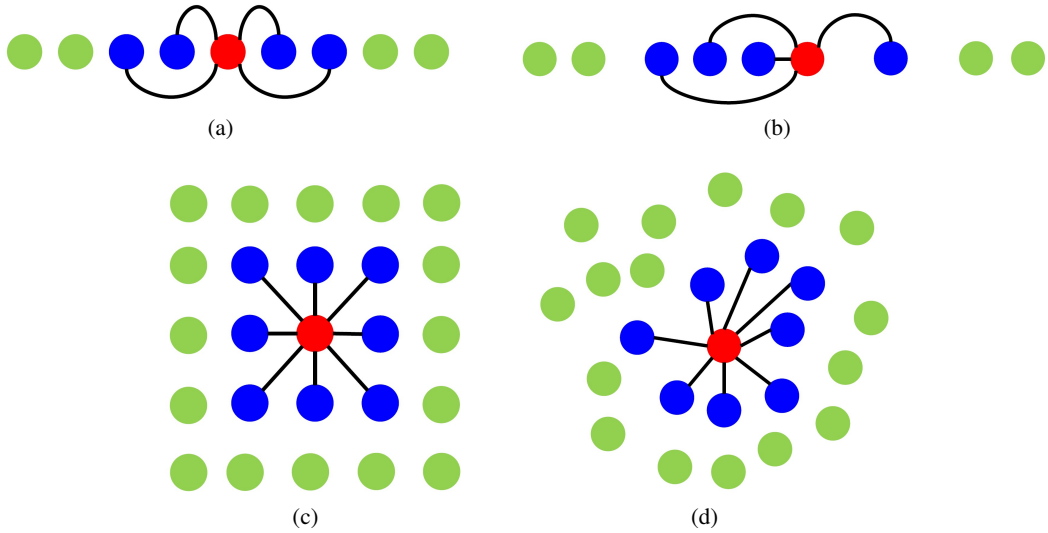


Figure 1: Graphic showing connections with 4 and 8 nearest neighbor nodes in a 1D and 2D problem setting respectively.

Once the graph network is designed using the method discussed above, we proceed with designing the architecture. This includes attention mechanism and different operations to be carried out within the network. Without loss of generality, let us consider the $i-$th node v_i in a 2D graph. As per the rule discussed above, v_i is connected to 8 neighbors. The edges connecting the nodes are denoted as $e_{i,k} := (v_i, v_k), k = 1, \ldots, 8$. For ease of understanding, we only focus on one edge $e_{i,j}$. We consider the spatial coordinates of the $i-$th and $j-$th nodes are $\boldsymbol{x}_i = [x_i, y_i]$ and $\boldsymbol{x}_j = [x_j, y_j]$, respectively. The graph architecture proposed in this work takes the relative difference between the spatial coordinates $\boldsymbol{\delta x}_{i,j} = (\boldsymbol{x}_i - \boldsymbol{x}_j)$ and the relative difference between the state variable $\boldsymbol{\delta u}_{i,j} = (\boldsymbol{u}_i - \boldsymbol{u}_j)$ as inputs. We consider a case where network is composed of 2 graph layers symbolised as $\mathscr{GN}^{(1)}$ and $\mathscr{GN}^{(2)}$. First graph layer provides $\boldsymbol{\delta x}_{i,j}$ as an input to a FNN $\mathscr{N}_{N_1}\left(\cdot; \boldsymbol{\theta}_N^{(1)}\right) : \boldsymbol{\delta x}_{i,j} \mapsto \boldsymbol{\gamma}_{i,j}$

$$\boldsymbol{\gamma}_{i,j} = \mathscr{N}_{N_1}\left(\boldsymbol{\delta x}_{i,j}; \boldsymbol{\theta}_N^{(1)}\right), \tag{14}$$

where $\boldsymbol{\gamma}_{i,j} = \left[\gamma_{ux}^{(i,j)}, \gamma_{uy}^{(i,j)}, \gamma_{vx}^{(i,j)}, \gamma_{vy}^{(i,j)}\right] \in \mathbb{R}^4$ is the vector of attention weights for message $\mathscr{M}_{i,j}$ from neighbor v_j. $\boldsymbol{\theta}_N^{(1)}$ represents the parameters of the FNN. The output of the FNN $\boldsymbol{\gamma}_{i,j}$ and the relative difference between the state variable $\boldsymbol{\delta u}_{i,j}$ are then provided as an input to an operator $\boldsymbol{H}$ and the operator outputs the gradients of the state variable $\nabla \boldsymbol{u}_i$ at node i,

$$\nabla \boldsymbol{u}_i = \sum_{j \in N_i} \boldsymbol{H}\left(\boldsymbol{\gamma}_{i,j}, \boldsymbol{\delta u}_{i,j}\right) / N_i, \tag{15}$$

where N_i represents the neighbors of the $i-$th node. In essence, we design the operator $\boldsymbol{H}$ to first replicate $\boldsymbol{\delta u}_{i,j}$ as $\tilde{\boldsymbol{u}}_{i,j} = [\boldsymbol{\delta u}_{i,j}, \boldsymbol{\delta u}_{i,j}]$ and then carry out a Hadamard product with the neural network output,

$$\boldsymbol{H}\left(\boldsymbol{\gamma}_{i,j}, \boldsymbol{\delta u}_{i,j}\right) = \mathscr{M}_{i,j} = \tilde{\boldsymbol{\gamma}}_{i,j} \odot \boldsymbol{\delta} \tilde{\boldsymbol{u}}_{i,j}, \tag{16}$$

where

$$\mathscr{M}_{i,j} = \begin{bmatrix} \left(u_x^{(i,j)}\right)_x & \left(u_y^{(i,j)}\right)_x \\ \left(u_x^{(i,j)}\right)_y & \left(u_y^{(i,j)}\right)_x \end{bmatrix}, \tilde{\boldsymbol{\gamma}}_{i,j} = \begin{bmatrix} \gamma_{ux}^{(i,j)} & \gamma_{vx}^{(i,j)} \\ \gamma_{uy}^{(i,j)} & \gamma_{vy}^{(i,j)} \end{bmatrix}, \text{ and } \boldsymbol{\delta} \tilde{\boldsymbol{u}}_{i,j} = \begin{bmatrix} \delta u_x^{(i,j)} & \delta u_x^{(i,j)} \\ \delta u_y^{(i,j)} & \delta u_y^{(i,j)} \end{bmatrix}. \tag{17}$$

$\odot$ in Eq. (16) denotes Hadamard product. Finally, we carry out a summation over the neighboring nodes to obtain the output of the first graph network layer. The basic premise here is that the first graph network layer, when trained, should output the gradients of the state vector $\nabla \boldsymbol{u}$. For sake of brevity, the overall operation carried out in the first layer of the graph network (at all nodes) is represented as

$$\nabla \boldsymbol{u} = \mathscr{GN}^{(1)}\left(\boldsymbol{x}, \boldsymbol{u}; \boldsymbol{\theta}_N^{(1)}\right), \tag{18}$$

with $\boldsymbol{\theta}_N^{(1)}$ being the parameters of the network.

The second graph network also functions in similar way as the first layer. Similar to the first graph network layer, we first provide the relative spatial coordinates as an input to a FNN $\mathscr{N}_{N_2}\left(\cdot;\boldsymbol{\theta}_N^{(2)}\right) : \boldsymbol{\delta x}_{i,j} \mapsto \boldsymbol{\beta}_{i,j}$

$$\boldsymbol{\beta}_{i,j} = \mathscr{N}_{N_2}\left(\boldsymbol{\delta x}_{i,j}; \boldsymbol{\theta}_N^{(2)}\right), \tag{19}$$

where $\boldsymbol{\beta}_{i,j} = [\beta_{uxx}^{(i,j)}, \beta_{uxy}^{(i,j)}, \beta_{uyx}^{(i,j)}, \beta_{uyy}^{(i,j)}, \beta_{vxx}^{(i,j)}, \beta_{vxy}^{(i,j)}, \beta_{vyx}^{(i,j)}, \beta_{vyy}^{(i,j)}] \in \mathbb{R}^8$ are the outputs from the FNN. Again, this step represents the attention mechanism with $\boldsymbol{\theta}_N^{(2)}$ representing the neural network parameters. The output from $\mathscr{N}_{N_2}\left(\cdot;\boldsymbol{\theta}_N^{(2)}\right)$, $\boldsymbol{\beta}_{i,j}$ and the relative difference between the output from the first graph network layer are provided as inputs to the operator $\boldsymbol{H}$ and the operator outputs the second derivative of the state variables,

$$\mathbb{H}(\boldsymbol{u}_i) = \sum_{j \in N_i} \boldsymbol{H}\left(\boldsymbol{\beta}_{i,j}, \boldsymbol{\delta}\nabla \boldsymbol{u}_{i,j}\right). \tag{20}$$

$\mathbb{H}(\boldsymbol{u}_i)$ in Eq. (20) consist of the Hessian of the two state-variables, $u_x^{(i)}$ and $u_y^{(i)}$ at node i

$$\mathbb{H}(\boldsymbol{u}_i) = \left[\mathbb{H}_s\left(u_x^{(i)}\right), \mathbb{H}_s\left(u_y^{(i)}\right)\right], \tag{21}$$

where $\mathbb{H}_s(\cdot)$ represents the Hessian operator. Accordingly,

$$\mathbb{H}_s\left(u_x^{(i)}\right) = \begin{bmatrix} \left(u_x^{(i)}\right)_{xx} & \left(u_x^{(i)}\right)_{xy} \\ \left(u_x^{(i)}\right)_{yx} & \left(u_x^{(i)}\right)_{yy} \end{bmatrix}, \quad \mathbb{H}_s\left(u_y^{(i)}\right) = \begin{bmatrix} \left(u_y^{(i)}\right)_{xx} & \left(u_y^{(i)}\right)_{xy} \\ \left(u_y^{(i)}\right)_{yx} & \left(u_y^{(i)}\right)_{yy} \end{bmatrix}. \tag{22}$$

and

$$\mathbb{H}(\boldsymbol{u}_i) = \begin{bmatrix} \left(u_x^{(i)}\right)_{xx} & \left(u_x^{(i)}\right)_{xy} & \left(u_y^{(i)}\right)_{xx} & \left(u_y^{(i)}\right)_{xy} \\ \left(u_x^{(i)}\right)_{yx} & \left(u_x^{(i)}\right)_{yy} & \left(u_y^{(i)}\right)_{yx} & \left(u_y^{(i)}\right)_{yy} \end{bmatrix} \tag{23}$$

$(\cdot)_{kl}$ in Eqs. (22) and (23) represents derivative with respect to variables k and l. $\boldsymbol{H}\left(\tilde{\boldsymbol{\beta}}_{i,j}, \delta\nabla\tilde{\boldsymbol{u}}_{i,j}\right)$ in Eq. (20) is computed as

$$\boldsymbol{H}\left(\tilde{\boldsymbol{\beta}}_{i,j}, \delta\nabla\tilde{\boldsymbol{u}}_{i,j}\right) = \tilde{\boldsymbol{\beta}}_{i,j} \odot \delta\nabla\tilde{\boldsymbol{u}}_{i,j}, \tag{24}$$

where $\tilde{\boldsymbol{\beta}}_{i,j} \in \mathbb{R}^{2\times 4}$ in Eq. (24) is a matrix formulated by using the output of $\mathscr{N}_{N_2}\left(\cdot;\boldsymbol{\theta}_N^{(2)}\right)$

$$\tilde{\boldsymbol{\beta}}_{i,j} = \begin{bmatrix} \beta_{uxx}^{(i,j)} & \beta_{uxy}^{(i,j)} & \beta_{vxx}^{(i,j)} & \beta_{vxy}^{(i,j)} \\ \beta_{uyx}^{(i,j)} & \beta_{uyy}^{(i,j)} & \beta_{vyx}^{(i,j)} & \beta_{vyy}^{(i,j)} \end{bmatrix}. \tag{25}$$

$\delta_\nabla \tilde{\boldsymbol{u}}_{i,j}$ in Eq. (24) is formulated by first creating a copy of each column of $\delta_\nabla \boldsymbol{u}_{i,j}$, which in turn is computed by using the output of the first graph network layer

$$\delta\nabla\tilde{\boldsymbol{u}}_{i,j} = \begin{bmatrix} \delta\left[\left(u_x^{(i,j)}\right)_x\right] & \delta\left[\left(u_x^{(i,j)}\right)_x\right] & \delta\left[\left(u_y^{(i,j)}\right)_x\right] & \delta\left[\left(u_y^{(i,j)}\right)_x\right] \\ \delta\left[\left(u_x^{(i,j)}\right)_y\right] & \delta\left[\left(u_x^{(i,j)}\right)_y\right] & \delta\left[\left(u_y^{(i,j)}\right)_y\right] & \delta\left[\left(u_y^{(i,j)}\right)_y\right] \end{bmatrix}, \tag{26}$$

where

$$\delta\left[\left(u_k^{(i,j)}\right)_l\right] = (u_k)_l^{(i)} - (u_k)_l^{(j)}, \;\; k,l = x,y. \tag{27}$$

i and j in superscript denotes the $i-$th and $j-$th nodes in the graph connected through edge $e_{i,j}$. For sake of brevity, we represent the overall operation carried out in second graph network layer as

$$\mathbb{H}(\boldsymbol{u}) = \mathscr{GN}^{(2)}\left(\boldsymbol{x}, \nabla\boldsymbol{u}; \boldsymbol{\theta}_N^{(2)}\right). \tag{28}$$

As the final piece of the puzzle, we concatenate the state vector $\boldsymbol{u}$ with the outputs of the two graph network layers, $\nabla \boldsymbol{u}$ and $\mathbb{H}(\boldsymbol{u})$ and pass it through a FNN to obtain the operator $\mathscr{N}_N(\cdot;\boldsymbol{\Theta})$ in Eq. (13). Mathematically, the overall operation being carried out inside the network can be represented as

$$\boldsymbol{y} = \mathscr{N}_N(\boldsymbol{h};\boldsymbol{\Theta}) = \mathscr{N}_N\left(\underbrace{\left[\boldsymbol{u}, \underbrace{\mathscr{GN}^{(1)}\left(\boldsymbol{x},\boldsymbol{u};\boldsymbol{\theta}_N^{(1)}\right)}_{\nabla \boldsymbol{u}}, \underbrace{\mathscr{GN}^{(2)}\left(\boldsymbol{x},\mathscr{GN}^{(1)}\left(\boldsymbol{x},\boldsymbol{u};\boldsymbol{\theta}_N^{(1)}\right);\boldsymbol{\theta}_N^{(2)}\right)}_{\mathbb{H}(\boldsymbol{u})}\right]}_{\boldsymbol{h}};\boldsymbol{\theta}_N^{(3)}\right), \tag{29}$$

with $\boldsymbol{\Theta} = \left[\boldsymbol{\theta}_N^{(1)}, \boldsymbol{\theta}_N^{(2)}, \boldsymbol{\theta}_N^{(3)}\right]$ and $\boldsymbol{y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3, \boldsymbol{y}_4]$. The neural network is supposed to learn the dependencies of $\boldsymbol{y}_i$ on $\boldsymbol{y}_{1:i-1}$ (see Eq. (13e)). Note that the graphical network includes the FNN used for inducing the attention mechanism. A schematic representation of the network architecture is shown in Fig. 2. $\boldsymbol{u}_t$ is computed by combining Eq. (29) with Eq. (13).

Remark 4: We note that the number of layers used in the graph network is in accordance with the highest order of the spatial derivative. In this paper, we have limited ourselves to PDEs having second order spatial derivatives only. Also dimension of output of $\mathscr{GN}^{(1)}$ and $\mathscr{GN}^{(2)}$ is dependent upon dimension of problem, we only discuss the dimensions used for a 2D problem.

Remark 5: Unlike the original work on GAT in [35], attention in GAT is introduced directly by using the nodal coordinates. This is possible because we are dealing with a physical domain where the nodal coordinates are available to us. Intuitively, GrADE assigns more "attention" to the edges that connects nearby nodes.

Remark 6: Although the mathematical expression in Eqs. (15) - (29) are expressed in matrix form, we have implemented it by expressing the same in vectorized form.

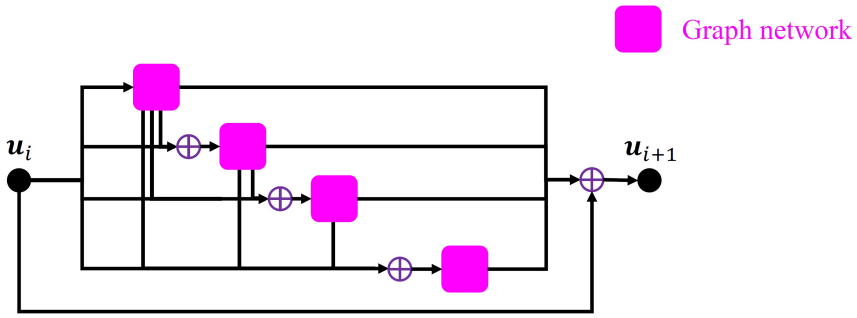
Remark 7: Although, we have mentioned gradient of state variables $\boldsymbol{u}$, $\nabla \boldsymbol{u}$ as output of the first graph network layer, this only holds when the network is trained. Similarly, once the network is trained, the second graph network layer yields the Hessian of the state variables. $\mathbb{H}(\boldsymbol{u})$.

4.2 Training

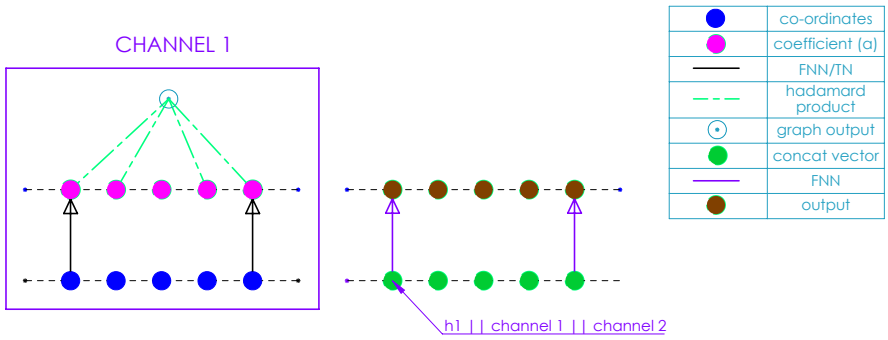
Having discussed the architecture of the proposed GrADE, we proceed to discuss the algorithm used for training the network. However, the proposed architecture blends GAT, FNN, and Neural ODE, each of which is trained differently. For instance, we generally use Message Passing (MP) algorithms for training graph networks. On the other hand, numerical integration schemes are used for training Neural ODE. Therefore, training GrADE naturally involves both MP and numerical integration, with MP being used in the spatial domain and numerical integration being used in the temporal domain.

Consider v_i to be the $i-$th node in the graph network. In MP, we update the state of v_i based on information from its neighbors N_i. The MP step can be divided into two steps: message aggregation and state update. In the message aggregation step, the messages received from all the nodes are aggregated into a single message. In the first graph layer of GrADE, the message aggregation step represents computing $\boldsymbol{H}(\gamma_{i,j}, \delta_{i,j})$ in Eq. (15) for each neighboring nodes followed by mean operation. Similarly, for the second graph layer of GrADE, the message aggregation step represents computing the mean in Eq. (20). The update step involve computing $\nabla \boldsymbol{u}$ from $\boldsymbol{u}$ in the first graph layer, and $\mathbb{H}(\boldsymbol{u})$ from $\nabla \boldsymbol{u}$ in the second graph layer. Once the updates for $\nabla \boldsymbol{u}$ and $\mathbb{H}(\boldsymbol{u})$ are available, we utilize the same in computation of the numerical integration. As already stated earlier, RK4 scheme is used in this paper for numerical integration.

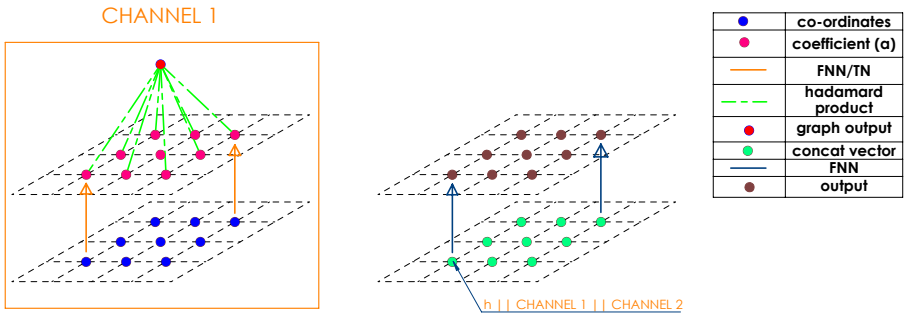
One major advantage of the proposed GrADE resides in the fact that time is not an explicit variable in the proposed framework; this allows the generalize the model better to future time-step. During the training phase, we allow the model to gradually explore the system and learn the network parameters. During the initial epochs, the proposed GrADE only explores a few steps starting from the initial condition. Slowly, as the model starts to learn, we allow GrADE to explore further time-steps. In practice, this is achieved by introducing a list variable τ_l that stores the number of time-steps GrADE is suppose to explore during each epoch. With this setup, GrADE is able to learn the dynamics that may be significantly different from the initial conditions and it neighbors. We also allow the learning rate η to vary with epoch by maintaining another list variable η_l. Overall, we implemented RK4 schme using the open-source library `torchdiffeq` [52]. For ease of understanding, an algorithm depicting the training procedure is shown in Algorithm 1.



(a) GrADE



(b) Graph network (1D)



(c) Graph network (2D)

Figure 2: Schematic representation of the proposed framework. (a) GrADE based on RK4-3/8 scheme. The magenta boxes represents $\boldsymbol{y}$ represented using Eq. (29). (b) Proposed graph attention for 1D problem, (c) Proposed graph attention for 2D problem.

5 Numerical implementation and results

We consider the well-known Burgers' equation for illustrating the performance of the proposed GrADE. Burgers' equation is a fundamental PDE occurring in various areas of applied mathematics, such as fluid mechanics, nonlinear acoustics, gas dynamics, and traffic flow. We solve Burger' equation in both 1D and 2D. For both cases, the simulation data is generated by using open-source FE solver, FeNICS [53].

Algorithm 1: Training GrADE

Inputs: $\mathscr{D} = \{\boldsymbol{x}, t_i, \boldsymbol{U}(\boldsymbol{x}, t_i)\}_{i=1}^{N_s}$.
Set Hyperparameter: Number of epochs N_e, time-step Δt, τ_l, and η_l.
Initialize: Neural network model: $\mathscr{N}_N(\cdot;\boldsymbol{\Theta})$; ▷ Eq. (29)
for *epoch* $= 0$ ***to*** N_e **do**
  $\tau \leftarrow \tau_l[epoch]$
  $\eta \leftarrow \eta_l[epoch]$
  Formulate $\boldsymbol{U}_t$ using $\boldsymbol{U}(\boldsymbol{x}, t_i)$ and $\tau_l[epoch]$; ▷ Target for current epoch
  for $t = 0$ ***to*** τ **do**
    $\boldsymbol{u}^{t+1} \leftarrow \boldsymbol{u}^t + \Delta t \times \mathscr{F}(\mathscr{N}_N(\cdot;\boldsymbol{\Theta}))$; ▷ Combination of Eqs. (29) and (13e)
    $\boldsymbol{U}_p[t] \leftarrow \boldsymbol{u}^{t+1}$; ▷ Store GrADE prediction
  end
  $\mathscr{L} = MSE(\boldsymbol{U}_p, \boldsymbol{U}_t)$; ▷ Calculate loss
  $\boldsymbol{\nabla w} \leftarrow$ Backprop($\mathscr{L}$)
  $\boldsymbol{w} \leftarrow \boldsymbol{w} - \eta \boldsymbol{\nabla w}$; ▷ Update weights
end
Output: Trained model $\mathscr{N}_N(\cdot;\boldsymbol{\Theta})$.

5.1 1D viscous Burgers' equation

First, we consider 1D viscous Burgers' equation with periodic boundary

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} - \nu\frac{\partial^2 u}{\partial x^2} = 0 \tag{30}$$

$$u(x=0,t) = u(x=L,t),\ x \in [0,L],\ t \in [0,T], \tag{31}$$

where u is the velocity and $\nu = 0.0025$ in viscosity. We consider random initial condition given by a Fourier series with random coefficients

$$u(x,t=0) = \frac{2w(x)}{\max_x |w(x)| + c}, \tag{32}$$

where

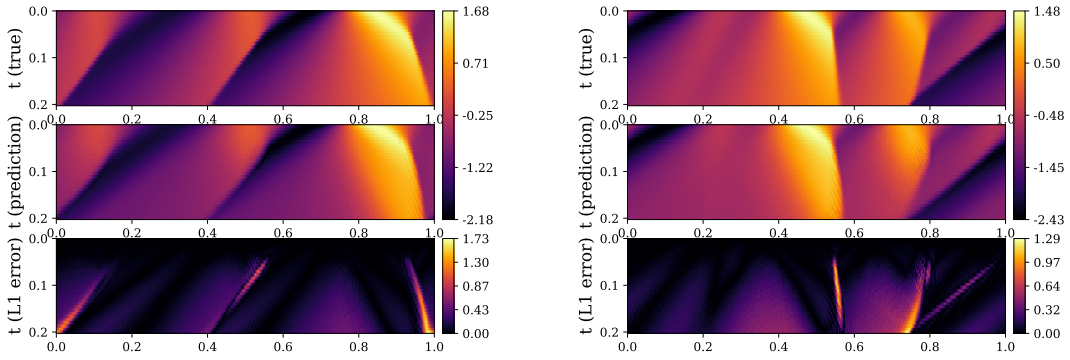
$$w(x) = a_0 + \sum_{l=1}^{N_l} a_l \sin(2l\pi x) + b_l \cos(2l\pi x). \tag{33}$$

In Eq. (33), $a_l, b_l \sim N(0,1)$ are drawn from standard Gaussian distribution and $c \sim \mathscr{U}(-1,1)$ is drawn from a uniform distribution. We have considered $L = 1$ in Eq. (31) and $N_l = 4$ in Eq. (33).

For generating data using FeNICS, we discretized the spatial domain into 512 points and use a time-step $\Delta t = 0.001$. For training and testing the proposed GrADE, we use $\Delta t = 0.007$. All the three FNNs present within the proposed GrADE are considered to be shallow nets with only hidden layer. We use LeakyReLU activation functions with a negative slope of 0.2 for all FNNs. For the two attention nets $\mathscr{N}_{N_1}\left(\cdot;\boldsymbol{\theta}_N^{(1)}\right)$ and $\mathscr{N}_{N_2}\left(\cdot;\boldsymbol{\theta}_N^{(2)}\right)$, the hidden layer has 32 neurons. As for the third network $\mathscr{N}_N\left(\cdot;\boldsymbol{\theta}_N^{(3)}\right)$, the hidden layer has 32 neurons. As we are dealing with a 1D problem here, all the three FNNs have only one output each. Overall the proposed GrADE has 387 parameters.

We trained the proposed GrADE using 120 samples of the initial condition and snapshots at four time instants only (i.e., last integration time index is 4). We use a learning rate of 0.07 and train the model for 201 epochs. For testing, we used 30 additional realizations of the random initial conditions. Fig. 3 shows the solutions of 1D Burgers' equation for two random initial condition (from the test set) obtained using FeNICS (first row) and GrADE (second row). While $x-$ axis in Fig. 3 represents the spatial domain, $y-$ axis represents the temporal domain. The third row in Fig. 3 represents the L1 error between the FeNICS and the GrADE results. We observe that results obtained using GrADE and FeNICS matches almost exactly. We note that the underlying dynamics is extremely complex due to formation of shocks. It is impressive that the proposed model trained with data at four temporal snapshots only (with $\Delta t - 0.007$) is able capture the temporal evolution of the system dynamics far beyond the training regime (up to 0.2s).

To illustrate the robustness of the proposed GrADE, we perform numerical experiments by varying the last integration time index and number of training scenarios provided to GrADE during training. For efficient training, different



(a) (b)

Figure 3: Figure depicting evolution of predicted velocities of 1D Burgers' equation with two initial conditions. First, second, third row shows FEM simulation, network prediction and L1 error, respectively.

Table 1: Hyper-parameters of GrADE for Burgers' 1D equation. * represents number of items in list, similar to python list notation.

lit train (index)	**learning rate list**	**training epochs**	**Training scenarios**
2	[0.07]*201	201	120
3	[0.07]*401	401	120
4	[0.05]*25 + [0.052]*25 + [0.054]*50 + [0.056]*301	401	120
5	[0.045]*25 + [0.048]*25 + [0.052]*50 + [0.054]*301	401	120

network hyperparameters have been used for different case. Details on the same is provided in Table 1. For comparing the accuracy in prediction, we compute the L2 error between the GrADE predicted results and true solution at each time-index as follows:

$$\varepsilon_j = \sum_{i=1}^{N_s} ||u_{p,j} - u_{t,j}||_2^2, \tag{34}$$

where N_s denotes the number of test samples, $u_{p,j}$ is the Grade predicted result at time-index j, and $u_{t,j}$ represents the target obtained using FeNICS. ε_j is the error at time-index j.

Fig. 4(a) shows results for Experiment 1 which compares prediction error with increasing time, for networks trained on different last integration time index. We notice that for last integration time index of two, three and four, the network has identical predictive capability. However, for last integration time index of 5, the result starts deviating beyond time-index 5, indicating over-fitting. As for computational time, the network trained with higher last integration time index takes more time to train. Fig 4(b) shows results for Experiment 2 which compares prediction error with increasing time for networks trained with different number of training graphs (training scenarios). We use a learning rate of 0.07 and last integration time index of 4 for all training sizes. As expected, we observe that the best result is obtained with 120 training scenarios and the worst with 30 scenarios. Results obtained with 60 and 90 scenarios are almost same.

Finally, we examine the output of the GAT present within the proposed GrADE. As stated earlier, once trained, the output of the first and second graph network layers should yield spatial derivatives u_x and u_{xx}, respectively. In Fig 5, we compare the outputs of the two graph network layers with the derivative obtained using central difference scheme. Excellent match between the two is observed.

5.2 2D coupled Burgers' equation

We consider the the 2D coupled Burgers' system. It has the same convective and diffusion form as the incompressible Navier-Stokes equations. It is an important model for understanding of various physical flows and problems, such as hydrodynamic turbulence, shock wave theory, wave processes in thermo-elastic medium, vorticity transport, dispersion in porous medium. Numerical solution of Burgers' equation is primary step towards when developing methods for

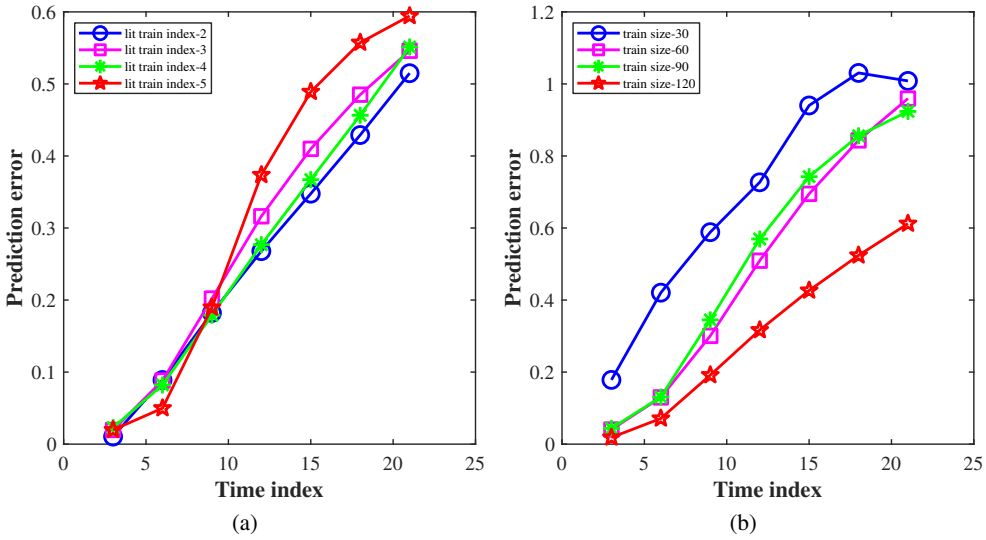


Figure 4: Burgers' 1d: (a) Prediction error with time for models trained with different last integration time index (LIT), where N_t is max time used during training (b) Prediction error with time for model trained with different number of training graph. Note that actual time is $Time\,index \times \Delta t$, where $\Delta t = 0.007$

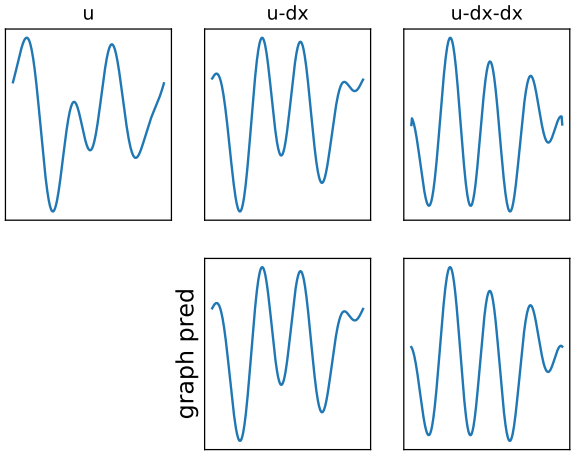


Figure 5: Caparison of output of two graph network layers used in model and derivatives of data computed using central differences.

complex flows. The governing equations for Burgers' equation takes the following form:

$$\boldsymbol{u}_t + \boldsymbol{u} \cdot \nabla \boldsymbol{u} - \nu \Delta \boldsymbol{u} = 0, \tag{35}$$

with periodic boundary condition

$$\begin{aligned} \boldsymbol{u}(x=0,y,t) &= \boldsymbol{u}(x=L,y,t), \\ \boldsymbol{u}(x,y=0,t) &= \boldsymbol{u}(x,y=L,t). \end{aligned} \tag{36}$$

Eq. (35) can be written in expanded form as

$$\begin{aligned}\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}-\nu(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2})=0\\ \frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}-\nu(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2})=0,\end{aligned} \tag{37}$$

where $\nu = 0.005$ is viscosity, u and v are the x and y components of velocity. We consider $\{x,y\} \in [0,1]$. Similar to the 1D case, the initial condition is defined using truncated Fourier series with random coefficients:

$$\boldsymbol{u}(x,y,t=0) = \frac{2\boldsymbol{w}(x,y)}{\max_{\{x,y\}}|\boldsymbol{w}(x,y)|}+\boldsymbol{c}, \tag{38}$$

where

$$\boldsymbol{w}(x,y) = \sum_{i=-L}^{N_l}\sum_{j=-L}^{L}\boldsymbol{a}_{ij}\sin(2\pi(ix+jy))+\boldsymbol{b}_{ij}\cos(2\pi(ix+jy)), \tag{39}$$

where $\boldsymbol{a}_{ij}, \boldsymbol{b}_{ij} \sim \boldsymbol{N}(0,\mathbf{I}_2)$, $L = 4$ and $\boldsymbol{c} \sim \mathscr{U}(-1,1) \in \mathbb{R}^2$. Some representative initial conditions generated using Eq. (38) are shown in Fig. 6.

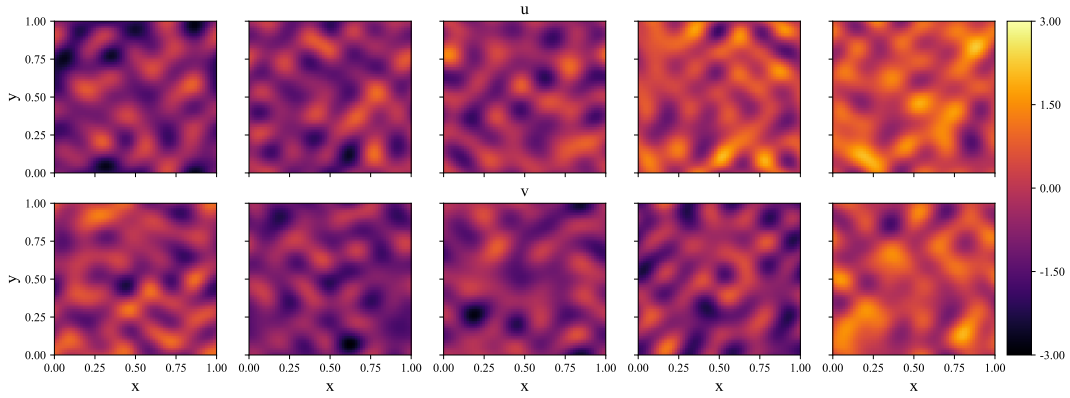
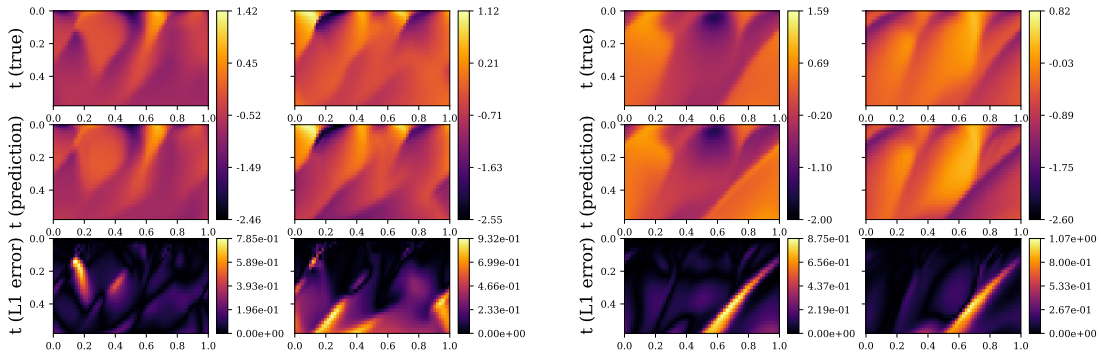


Figure 6: Randomly generated initial condition for x and y velocity-components using truncated Fourier series.

Similar to the 1D case, we use FeNICS to generate the training data. We discretize the spatial domain in FeNICS into 64×64 grid and use a time-step of 0.005. For the FNNs present within the proposed GrADE, we consider shallow nets with only one hidden layer. For the two attention nets $\mathscr{N}_{N_1}\left(\cdot;\boldsymbol{\theta}_N^{(1)}\right)$ and $\mathscr{N}_{N_2}\left(\cdot;\boldsymbol{\theta}_N^{(2)}\right)$, the hidden layer has 32 neurons. The hidden layer of the third network $\mathscr{N}_N\left(\cdot;\boldsymbol{\theta}_N^{(3)}\right)$ has 64 neurons. Overall, the proposed GrADE has 2446 trainable parameters.

We trained the proposed GrADE using 120 samples of the initial condition and snapshots at three time instants only (i.e., last integration time index is 3). We allowed the learning rate to vary with number of epochs and trained the model for 501 epochs. For testing, we generated 20 additional realizations of the random initial conditions. Fig. 7 shows the the results corresponding to two initial conditions from the test dataset obtained using FeNICS and the proposed GrADE. The first and second rows depict the velocities (slices along x and y axes) obtained using FeNICS and the proposed approach respectively. Reasonable match among the results is observed. The third column shows the L1 error. Fig. 8 shows the velocities at different time-steps obtained using FeNICS and the proposed approach. Note that predicting the velocities for the 2D case as well is extremely difficult because of the formation of shocks. The fact that the proposed GrADE trained with observations at only three snapshots (with $\Delta t = 0.02$) is able to provide reasonably accurate results is really impressive.

To understand the influence of last integration time index used during training, we perform case study by varying the last integration time index. The hyperparameter setting for all the cases are shown in Table 2. The results obtained are shown in Fig. 9. Unlike the 1D Burgers' equation where the error was almost similar for all the cases, error is least when trained with last integration time index of 3. This is probably because the loss-function becomes extremely



(a) (b)

Figure 7: Figure depicting domain slice evolution of predicted x and y velocities of 2D coupled Burgers' equation with two initial conditions. Last row depicts the L1 error.

complex on increasing the last integration time index during training beyond 3. One way to address this issue is to use depth refinement. The idea is to alter depth of network during training. We use a smaller last integration time index during the initial training; however, as the training progresses, we start increasing the depth of the model. We recall that depth in GrADE refers to the number of time integration steps. This method enable us to train networks with more depth. The results corresponding to depth refinement are shown in Fig. 11. We start with a depth of 2 and gradually increased it till depth 4. Note that extra care is necessary with the depth refinement framework. The hyperparameters used are shown in Table 3. Results are compared with those obtained using a constant depth of 4. We observe that the computational time needed is less and the accuracy of the model is better for the depth refinement framework.

Table 2: Hyper-parameters of GrADE for Burgers' 2D equation for Experiment 2. * represents number of items in list, similar to python list notation.

lit train	learning rate list at each training epoch	training epochs	training initial conditions
2	[0.055]*200 + [0.053]*100 + [0.05]*101	401	90
3	[0.055]*200 + [0.054]*100 + [0.03]*101	401	90
4	[0.045]*400 + [0.044]*301	701	90

Table 3: Hyper-parameters of GrADE for Burgers' 2D equation for Experiment 3. * represents number of items in list, similar to python list notation. Model 1 is constant depth and model 2 is with depth refinement

model type	lit train	learning rate list at each training epoch	training epochs	training initial conditions
1	4	[0.045]*701	701	90
2	[2] * 200 + [3] * 300 + [4] * 201	[0.06] * 200 + [0.022] * 25 + [0.024] * 25 + [0.032] * 50 + [0.04] * 200 + [0.015] * 25 + [0.018] * 25 + [0.022] * 25 + [0.032] * 25 + [0.04] * 101	701	90

Next we concentrate on the role of attention model within the proposed framework. In this work, we have used FNN for computing the attention weights $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. An alternative to this is to use Taylor net. SpiderConv proposed in [54] uses Taylor net in GNN for classification and segmentation tasks. It can be formulated as

$$\boldsymbol{\gamma}_{i,j} = \sum_{k=0}^{Q} w_k * p_k(\boldsymbol{\delta x}_{i,j}), \tag{40}$$

where p_k is element of $p \in \boldsymbol{\pi}_m(\mathbb{R}^2)$, m is degree of polynomials, Q is number of monomials and w_k are trainable weights of networks. In Fig. 11, we present a comparative assessment between model accuracy when using FNN and Taylor net. We use $m = 3, Q = 10$ in Eq. (40). We observe that the results obtained using the FNN is slightly more accurate as

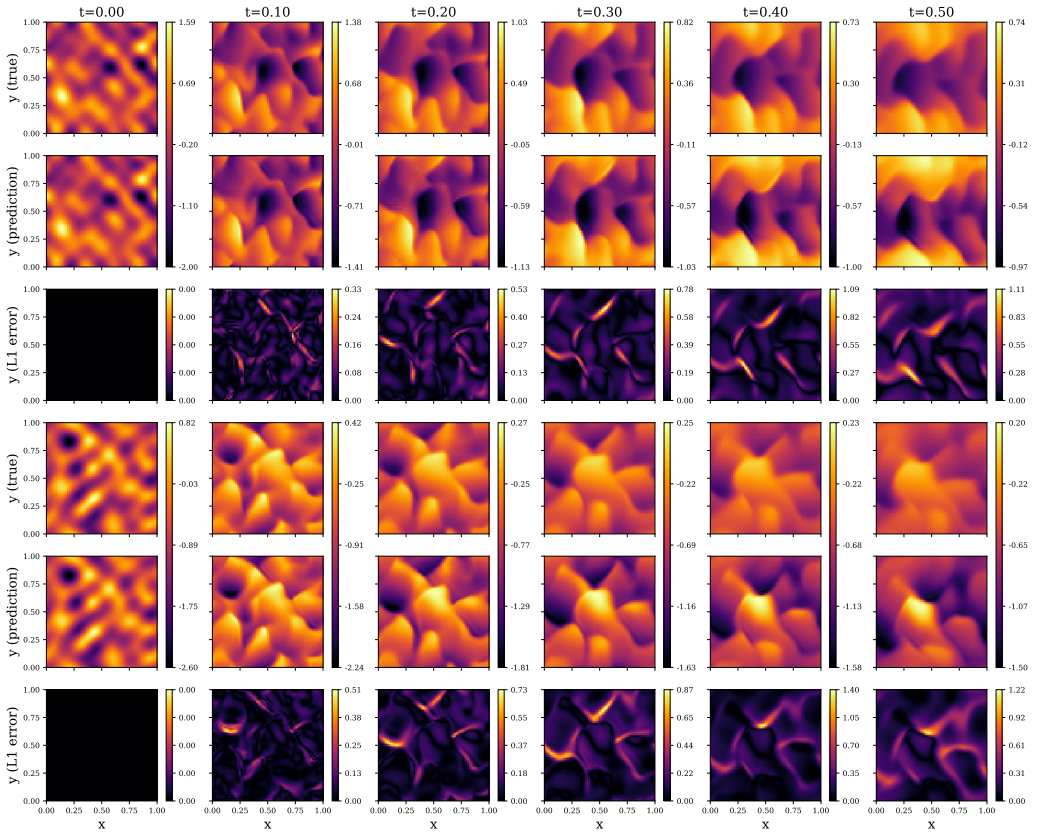


Figure 8: Prediction of x and y velocities of 2D coupled Burgers' equation at different time steps. Top to bottom, three rows shows x-velocity, y-velocity, L1 error related plots, respectively.

compared to Taylor net. Recall that the output of the first graph network layer is supposed to yield the first derivative and those of the second layer is supposed to yield the second derivative. To validate the same, we plot the output of the two graph network layers in Fig. 12. Results obtained using central difference are also shown. A reasonably good match between the output of the graph network layers and those obtained using central difference is observed. This illustrates that the graph is able to capture the spatial derivatives.

6 Conclusions

In this work, we have presented a novel data-driven framework for solving time-dependent nonlinear partial differential equations (PDE). The proposed approach is referred to as Graph Attention PDE or GrADE couples Feed-forward Neural Networks (FNN), Graph Attention (GAT), and Neural Ordinary Differential Equation (Neural ODE). The key idea is to use GAT to model the spatial domain an Neural ODE to model the temporal domain. FNNs are used for modeling the attention mechanism within the GAT network. GAT ensures that the problem at hand is computationally tractable as a node in the graph is only connected to its neighbors. Neural ODE, on the other hand, results in constant memory cost and allows trading of numerical precision for speed. While different numerical time-integration schemes can be used within the proposed framework, we have use forth order Runge Kutta method in this work. We also proposed depth refinement as an effective technique for training the proposed architecture in lesser time time with better accuracy.

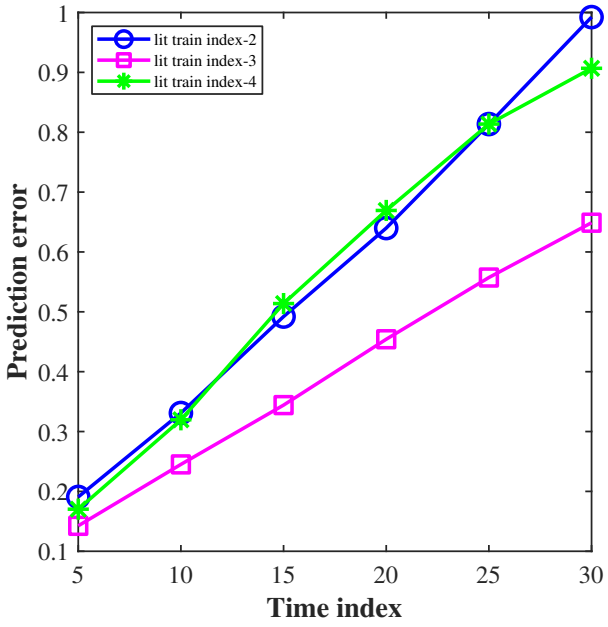


Figure 9: Prediction error with increasing time for model trained with different number of training graph

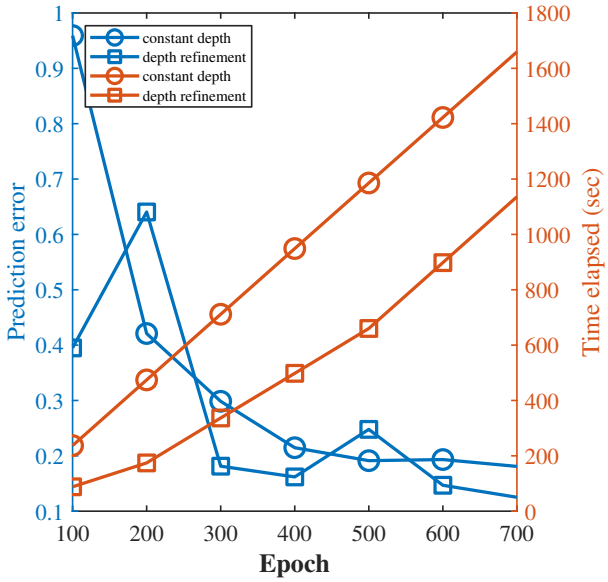


Figure 10: Plot depicts prediction error and time elapsed at different training epochs during training GrADE with constant depth versus with depth refinement.

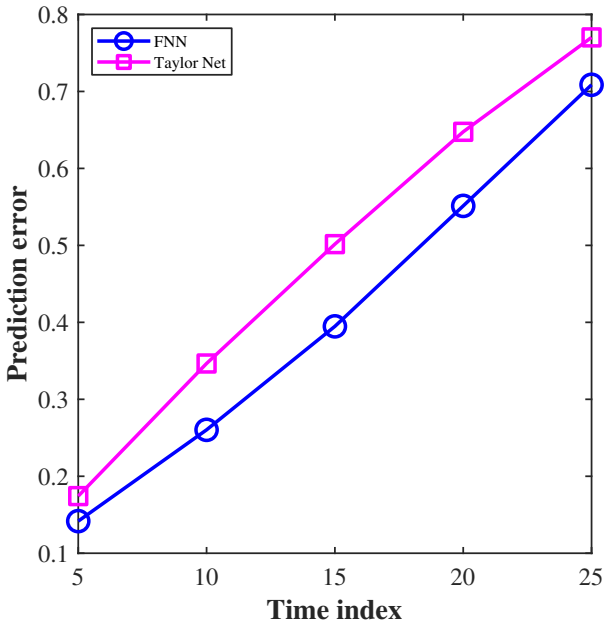


Figure 11: Prediction error with time for model using a FNN and Taylor net for $\mathcal{N}_{N_1}$ and $\mathcal{N}_{N_2}$ in eq. 14 and 19

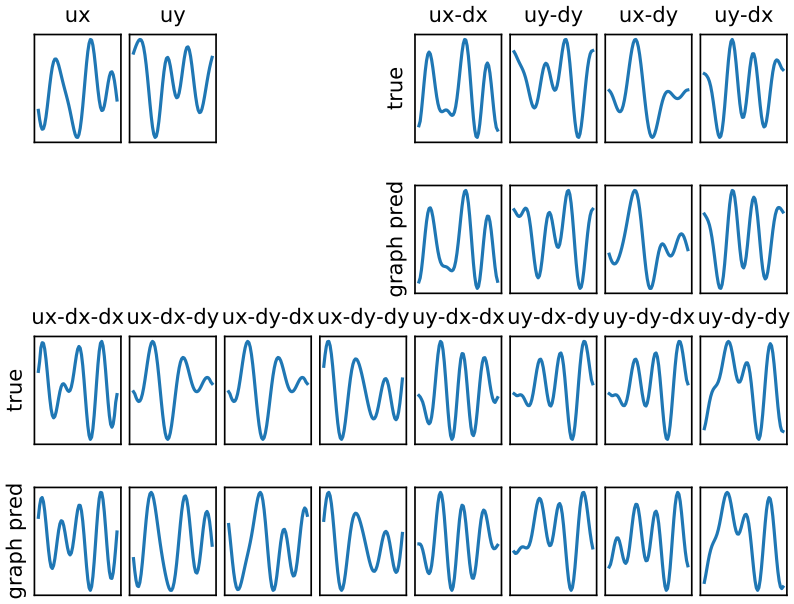


Figure 12: Caparison of output of two graph network used in model and derivatives of data computed using central differences.

We solve Burgers' equation to illustrate the performance of the proposed approach. Both 1D and 2D Burgers' equation has been solved. Results obtained have been benchmarked against those obtained using finite element solver. We observe that the proposed approach is able to provide accurate solution by using a larger time-step and snapshots of data at only two time-instants (referred as last integration time index). Case studies by varying last integration time

index and amount of training data showcase the robustness of the proposed approach. We illustrated that the graph network layers are able to accurately capture the spatial derivatives of the state variables. We also showed that using depth refinement training strategy can help reduce training time for the network and increase its accuracy.

Acknowledgements: SC acknowledges the financial support received in form of seed grant from IIT Delhi.

References

[1] Olgierd Cecil Zienkiewicz, Robert Leroy Taylor, Perumal Nithiarasu, and JZ Zhu. *The finite element method*, volume 3. McGraw-hill London, 1977.

[2] Fadl Moukalled, L Mangani, Marwan Darwish, et al. *The finite volume method in computational fluid dynamics*, volume 113. Springer, 2016.

[3] Tadeusz Liszka and Janusz Orkisz. The finite difference method at arbitrary irregular grids and its application in applied mechanics. *Computers & Structures*, 11(1-2):83–95, 1980.

[4] Mohammad H Aliabadi. *The boundary element method, volume 2: applications in solids and structures*, volume 2. John Wiley & Sons, 2002.

[5] Souvik Chakraborty and Nicholas Zabaras. Efficient data-driven reduced-order models for high-dimensional multiscale dynamical systems. *Computer Physics Communications*, 230:70–88, 2018. ISSN 0010-4655. doi: https://doi.org/10.1016/j.cpc.2018.04.007.

[6] Han Gao, Jian-Xun Wang, and Matthew J. Zahr. Non-intrusive model reduction of large-scale, nonlinear dynamical systems using deep learning. *Physica D: Nonlinear Phenomena*, 412:132614, 2020. ISSN 0167-2789. doi: https://doi.org/10.1016/j.physd.2020.132614.

[7] Dongbin Xiu and George Em Karniadakis. The wiener–askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing*, 24(2):619–644, 2002. doi: 10.1137/S1064827501387826.

[8] Ilias Bilionis, Nicholas Zabaras, Bledar A. Konomi, and Guang Lin. Multi-output separable gaussian process: Towards an efficient, fully bayesian paradigm for uncertainty quantification. *Journal of Computational Physics*, 241:212–239, 2013. ISSN 0021-9991. doi: https://doi.org/10.1016/j.jcp.2013.01.011.

[9] Steven Atkinson and Nicholas Zabaras. Structured bayesian gaussian process latent variable model: Applications to data-driven dimensionality reduction and high-dimensional inversion. *Journal of Computational Physics*, 383: 166–195, 2019. ISSN 0021-9991. doi: https://doi.org/10.1016/j.jcp.2018.12.037.

[10] Michael Schmidt and Hod Lipson. Distilling free-form natural laws from experimental data. *Science*, 324(5923): 81–85, 2009. ISSN 0036-8075. doi: 10.1126/science.1165893.

[11] Josh Bongard and Hod Lipson. Automated reverse engineering of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 104(24):9943–9948, 2007. ISSN 0027-8424. doi: 10.1073/pnas.0609476104.

[12] Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15): 3932–3937, 2016. ISSN 0027-8424. doi: 10.1073/pnas.1517384113.

[13] Maziar Raissi and George Em Karniadakis. Hidden physics models: Machine learning of nonlinear partial differential equations. *Journal of Computational Physics*, 357:125–141, 2018. ISSN 0021-9991. doi: https://doi.org/10.1016/j.jcp.2017.11.039.

[14] Ravi G. Patel and Olivier Desjardins. Nonlinear integro-differential operator regression with neural networks, 2018.

[15] Nicholas Geneva and Nicholas Zabaras. Transformers for modeling physical systems, 2021.

[16] Nicholas Zabaras Nicholas Geneva. Multi-fidelity generative deep learning turbulent flows. *Foundations of Data Science*, 2(4):391–428, 2020.

[17] Romit Maulik, Romain Egele, Bethany Lusch, and Prasanna Balaprakash. Recurrent neural network architecture search for geophysical emulation. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, SC '20. IEEE Press, 2020. ISBN 9781728199986.

[18] Nicholas Geneva and Nicholas Zabaras. Modeling the dynamics of pde systems with physics-constrained deep auto-regressive networks. *Journal of Computational Physics*, 403:109056, 2020. ISSN 0021-9991. doi: https://doi.org/10.1016/j.jcp.2019.109056.

[19] Kailiang Wu and Dongbin Xiu. Numerical aspects for approximating governing equations using data. *Journal of Computational Physics*, 384:200–221, 2019. ISSN 0021-9991. doi: https://doi.org/10.1016/j.jcp.2019.01.030.

[20] I.E. Lagaris, A. Likas, and D.I. Fotiadis. Artificial neural networks for solving ordinary and partial differential equations. *IEEE Transactions on Neural Networks*, 9(5):987–1000, 1998. doi: 10.1109/72.712178.

[21] M. Raissi, P. Perdikaris, and G.E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019. ISSN 0021-9991. doi: https://doi.org/10.1016/j.jcp.2018.10.045.

[22] Henning Wessels, Christian Weißenfels, and Peter Wriggers. The neural particle method–an updated lagrangian physics informed neural network for computational fluid dynamics. *Computer Methods in Applied Mechanics and Engineering*, 368:113127, 2020.

[23] Shengze Cai, Zhicheng Wang, Sifan Wang, Paris Perdikaris, and George Em Karniadakis. Physics-informed neural networks for heat transfer problems. *Journal of Heat Transfer*, 143(6):060801, 2021.

[24] Somdatta Goswami, Cosmin Anitescu, Souvik Chakraborty, and Timon Rabczuk. Transfer learning enhanced physics informed neural network for phase-field modeling of fracture. *Theoretical and Applied Fracture Mechanics*, 106:102447, 2020.

[25] Souvik Chakraborty. Simulation free reliability analysis: A physics-informed deep learning based approach. *arXiv preprint arXiv:2005.01302*, 2020.

[26] Minliang Liu, Liang Liang, and Wei Sun. A generic physics-informed neural network-based constitutive model for soft biological tissues. *Computer methods in applied mechanics and engineering*, 372:113402, 2020.

[27] Yinhao Zhu, Nicholas Zabaras, Phaedon-Stelios Koutsourelakis, and Paris Perdikaris. Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data. *Journal of Computational Physics*, 394:56–81, 2019.

[28] Souvik Chakraborty. Transfer learning based multi-fidelity physics informed deep neural network. *Journal of Computational Physics*, 426:109942, 2021.

[29] Xuhui Meng and George Em Karniadakis. A composite neural network that learns from multi-fidelity data: Application to function approximation and inverse pde problems. *Journal of Computational Physics*, 401:109020, 2020.

[30] Sasha Targ, Diogo Almeida, and Kevin Lyman. Resnet in resnet: Generalizing residual architectures. *arXiv preprint arXiv:1603.08029*, 2016.

[31] Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David Duvenaud. Neural ordinary differential equations. *arXiv preprint arXiv:1806.07366*, 2018.

[32] Emilien Dupont, Arnaud Doucet, and Yee Whye Teh. Augmented neural odes. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

[33] Juntang Zhuang, Nicha Dvornek, Xiaoxiao Li, Sekhar Tatikonda, Xenophon Papademetris, and James Duncan. Adaptive checkpoint adjoint method for gradient estimation in neural ode, 2020.

[34] Qiqi Wang, Parviz Moin, and Gianluca Iaccarino. Minimal repetition dynamic checkpointing algorithm for unsteady adjoint calculation. *SIAM Journal on Scientific Computing*, 31(4):2549–2567, 2009. doi: 10.1137/080727890.

[35] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks, 2018.

[36] Tong Qin, Kailiang Wu, and Dongbin Xiu. Data driven governing equations approximation using deep neural networks. *Journal of Computational Physics*, 395:620–635, 2019.

[37] Allan Pinkus. Approximation theory of the mlp model in neural networks. *Acta Numerica*, 8:143–195, 1999. doi: 10.1017/S0962492900002919.

[38] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*, 2014.

[39] Minh-Thang Luong, Hieu Pham, and Christopher D Manning. Effective approaches to attention-based neural machine translation. *arXiv preprint arXiv:1508.04025*, 2015.

[40] Heliang Zheng, Jianlong Fu, Tao Mei, and Jiebo Luo. Learning multi-attention convolutional neural network for fine-grained image recognition. In *Proceedings of the IEEE international conference on computer vision*, pages 5209–5217, 2017.

[41] Ruben Rodriguez-Torrado, Pablo Ruiz, Luis Cueto-Felgueroso, Michael Cerny Green, Tyler Friesen, Sebastien Matringe, and Julian Togelius. Physics-informed attention-based neural network for solving non-linear partial differential equations. *arXiv preprint arXiv:2105.07898*, 2021.

[42] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008, 2017.

[43] Kevin P. Murphy. *Probabilistic Machine Learning: An introduction*. MIT Press, 2022.

[44] Mikael Henaff, Joan Bruna, and Yann LeCun. Deep convolutional networks on graph-structured data, 2015.

[45] Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst. Convolutional neural networks on graphs with fast localized spectral filtering. In D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc., 2016.

[46] Ron Levie, Federico Monti, Xavier Bresson, and Michael M. Bronstein. Cayleynets: Graph convolutional neural networks with complex rational spectral filters. *IEEE Transactions on Signal Processing*, 67(1):97–109, 2019. doi: 10.1109/TSP.2018.2879624.

[47] Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. The graph neural network model. *IEEE Transactions on Neural Networks*, 20(1):61–80, 2009. doi: 10.1109/TNN.2008.2005605.

[48] Will Hamilton, Zhitao Ying, and Jure Leskovec. Inductive representation learning on large graphs. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

[49] F. Monti, D. Boscaini, J. Masci, E. Rodola, J. Svoboda, and M. M. Bronstein. Geometric deep learning on graphs and manifolds using mixture model cnns. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5425–5434, Los Alamitos, CA, USA, jul 2017. IEEE Computer Society. doi: 10.1109/CVPR.2017.576.

[50] Alejandro F Queiruga, N Benjamin Erichson, Dane Taylor, and Michael W Mahoney. Continuous-in-depth neural networks. *arXiv preprint arXiv:2008.02389*, 2020.

[51] Leif E Peterson. K-nearest neighbor. *Scholarpedia*, 4(2):1883, 2009.

[52] Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David Duvenaud. Neural ordinary differential equations. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 6572–6583, Red Hook, NY, USA, 2018. Curran Associates Inc.

[53] Martin Alnæs, Jan Blechta, Johan Hake, August Johansson, Benjamin Kehlet, Anders Logg, Chris Richardson, Johannes Ring, Marie E Rognes, and Garth N Wells. The fenics project version 1.5. *Archive of Numerical Software*, 3(100), 2015.

[54] Yifan Xu, Tianqi Fan, Mingye Xu, Long Zeng, and Yu Qiao. Spidercnn: Deep learning on point sets with parameterized convolutional filters. September 2018.

Current Psychology
https://doi.org/10.1007/s12144-021-02169-6

Gratitude resentment and appreciation scale (GRAT-16): analyzing psychometrics properties in the Indian context

Naval Garg[1] · Mehak[2]

Accepted: 28 July 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract
GRAT-16 is one of the most commonly used scales that measure gratitude. This scale was designed and majorly used in the western context. The present paper examined the factorial solution, reliability, and validity of the scale in Indian settings. For psychometric analysis of GRAT-16, two studies were conducted. The first study examined the factorial validity of the three-dimensional scale of gratitude (GRAT-16) in the Indian context. The first study was based on a sample of 530 students that consisted of 51.32% male and 48.68% female respondents. These students were enrolled in commerce (68.11%), science (20.57%), and humanities (11.32%) courses and have a mean age of 21.19 years. Study 1 had two sub-samples. The first subsample ($n = 260$) was used for exploratory factor analysis using varimax rotation. The factorial solution reported three factors that explain a 54% variance. The second subsample ($n = 270$) was used for confirmatory factor analysis to substantiate a three-factor scale. Here, one statement was deleted for low regression weight. It reduced GRAT-16 to GRAT-15 for Indian respondents with appreciable goodness of fit indices. The second study examined the reliability and validity of the Indian GRAT-15. Study 2 was based on a sample of 519 students, with a mean age of 21.65 years. This sample had 288 (53.56%) male participants and 241 (53.56%) female respondents. Also, 63.2%, 26.37%, and 10.31% of the participants belonged to commerce, science, and humanities streams. This second study concluded good internal consistency reliability with acceptable values of Cronbach's alpha and composite reliability. The appropriate values of Average Variance Explained (AVE) confirmed convergent and discriminant validity of Indian GRAT-15. The criterion validity of the Indian GRAT-15 was also assessed. A significant positive correlation with the Satisfaction with Life Scale and a negative association with Kessler Psychological Distress Scale concluded that GRAT-15 might be a valuable tool to measure gratitude for the Indian respondents.

Keywords Gratitude · GRAT-16 · Reliability · Validity

Introduction

Previous researchers refer to gratitude with several connotations like an emotion, a trait, an affective characteristic, an inherent virtue, a habit, and a mood (Emmons & McCullough, 2003). In the words of Watkins et al. (2014), "in various contexts, gratitude can refer to an emotional state, an emotional expression, a character trait or even a virtue" (p. 13). These multiple connotations lead to different conceptualizations of gratitude. The emotion of gratitude is experienced when one realizes that someone else has done something good for them (Watkins et al., 2014). As an emotion, it acts as a moral barometer, moral motivator, and moral reinforcer (McCullough et al., 2001). As an affective trait, gratitude is viewed as a "general tendency to recognize and respond with grateful emotions to the roles of other people's benevolence in the positive experiences and outcomes that one obtains" (McCullough et al., 2002, p. 112). As a habit, it is a general tendency to appreciate and enjoy good in life. As a mood, it is a transitory and short-lived state of being thankful for the benefactor. Apart from these conceptualizations, the researchers have elaborated on different dimensions of gratitude. Watkins et al. (2013) discussed three dimensions of gratitude- lack of sense of deprivation (LOSD), simple appreciation (SA), and appreciation for others (AO). Wood et al. (2010) suggested eight conceptions of the trait gratitude- individual differences, appreciation, have focus, awe, behavior, present moment, transient life, and favourable social

✉ Naval Garg
naval.garg@dtu.ac.in

[1] University School of Management and Entrepreneurship, Delhi Technological University, Delhi, India

[2] Indian Institute of Science, Bengaluru, India

Published online: 09 August 2021

comparisons. The authors defined gratitude as a higher-order construct that includes all the eight characteristics mentioned above.

Recent researchers have focused mainly on the linkage of gratitude with personal, psychological, social, and organizational variables. At the workplace, it positively influences employee well-being (McCullough et al., 2001), leadership (Garg, 2020), organizational citizenship behaviour (Fehr et al., 2017; Garg et al., 2019), efficiency and loyalty of the employees (Grant & Wrzesniewski, 2010), and prosocial work behaviors (Grant & Gino, 2010). It has a negative impact on employee turnover (Garg et al., 2021), resentment and envy, workplace incivility and harassment (Sarkar & Garg, 2021), and burnout (Chan, 2011). Other studies report that grateful people have better physical health (Krause & Hayward, 2014), sleep (Mills et al., 2015), emotional well-being (Millstein et al., 2016), happiness (McCullough et al., 2002), optimism (Hill & Allemand, 2011), hope (Sarkar & Garg, 2020), and positive affect (Watkins et al., 2003). Socially, gratitude is poised to develop long-lasting relationships among people. Researchers refer to it as "booster shots" for maintaining long-run associations. It promotes prosocial behaviour (Michie, 2009), trust (Drążkowski et al., 2017), mutual understanding, and respect (Algoe & Zhaoyang, 2015), and affection among people. In a nutshell, researchers recommend that gratitude enhances people's well-being. Emmons and Mishra (2011) provide ten pathways through which gratitude leads to well-being: 1) promote physical health, 2) help in goal achievement, 3) develop spirituality, 4) reduce feverishness for worldly pursuits, 5) enhance positive self-evaluations, 6) build and strengthen social relationships, 7) provide coping skills for challenging and stressful situations, 8) dilute the feeling of jealousy and envy, 9) enhance positive memories, and 10) motivate behaviours and actions for the benefit of others.

These benefits are theoretically derived from different theories of gratitude. First, the 'broaden-and-build theory of positive emotions' elaborates that gratitude widens a person's array of thoughts and actions. This broadening of one's outlook leads to the development of personal, psychological, and social resources for an individual (Fredrickson, 2004). The psychological resources include hope, trust, optimism, self-efficacy, vitality, and resilience (Garg & Sarkar, 2020). Social resources enlist friendships, long-term associations, social bonds, group cohesiveness, and team spirit. Second, Algoe's (2012) 'find-remind-bind' theory also explains gratitude's social and psychological utility. Gratitude makes one realize the goodness in self and others. This realization is reinforced through repeated reminders. And ultimately, gratitude blesses one with the perennial joy of goodness. Other theories like 'amplification theory,' 'social cognitive theory,' and 'social exchange theory' also provide a theoretical foundation to gratitude studies.

Acknowledging the countless benefits of gratitude, several instruments are developed by researchers to measure gratitude. First, the Gratitude Questionnaire (GQ-6) is a unidimensional scale that measures the disposition to experience gratitude. McCullough et al. (2002) reported excellent internal reliability, with Cronbach's alpha values between 0.82 to 0.87. Second, Gratitude Adjective Checklist (GAC), a three-item measure, is based on the average of affect adjectives: grateful, thankful, and appreciative. It is a flexible scale in which time may be varied as per the requirement. Respondents may be asked about their gratefulness in different time zones like right now, yesterday, past week, or the past few weeks. Third, Transpersonal Gratitude Scale (TGS) measures four dimensions of gratitude: an expression of gratitude, spiritual connection, the value of gratitude, and transcendent gratitude. It is a 16-item scale with good internal reliability and constructs validity. Fourth, the Multi-Component Gratitude Measure (MCGM) accesses four components of gratitude: understanding of gratitude, grateful emotions, attitudes toward gratitude, and gratitude-related behaviours. And lastly, Gratitude Resentment and Appreciation Scale (GRAT) is a 44-item scale with good internal consistency, factorial validity, construct validity, and temporal stability (Watkins et al., 2003). The GRAT-16 is a short-form and comprises sixteen items of GRAT-44 (Thomas & Watkins, 2003).

Although several scales are available to comprehend gratitude, these scales require validation in different cultural contexts due to gratitude's cultural sensitivity. Morgan et al. (2014) argue that "gratitude may contain a common core with culturally ubiquitous features, in addition to socially constructed elements that change depending on the culture being studied." Experience and expression of gratitude may vary with culture and nationality (Sommers & Kosmitzki, 1988). Tudge et al. (2015) report culture-based variability of gratitude among American, Brazilian, Chinese, and Russian children. In this background, the present study investigates the adaptability of GRAT-16 in the Indian context.

Gratitude Resentment and Appreciation Scale (GRAT-16)

The GRAT-16 or S-GRAT (Short-GRAT) is a short form of the 44-item Gratitude Resentment and Appreciation Scale. Watkins et al. (2003) initially drafted 53 items to tap four essential domains of sense of abundance, simple appreciation, appreciation for others, and importance of gratitude expression. Nine items were deleted owing to low item-total correlations. It led to the acceptance of a scale of 44 statements. Later on, Thomas and Watkins (2003) suggested a short form of scale for ease of data collection, resource conservation, and convenience of researchers. The short-GRAT scale comprised sixteen statements that measure three subscales of gratitude:

lack of sense of deprivation or sense of abundance (six items), simple appreciation (six items), and appreciation for others (four items). The statements are rated on a nine-point scale ranging from 1 (strongly disagree) to 9 (strongly agree). This GRAT-16 scale demonstrated satisfactory internal consistency (Cronbach's alpha = 0.92) in a sample of undergraduate students. Thomas and Watkins (2003) observed that GRAT-16 was positively correlated with life satisfaction (r = 0.65) and positive affectivity (r = 0.47). They also reported negative relationship of GRAT-16 with current depression (r = 0.43) and negative affectivity (r = 0.37). Short GRAT was as reliable and valid as GRAT (Thomas & Watkins, 2003). Five statements were reverse coded to reduce acquiescence bias. The total score of GRAT-16 ranged from 16 to 144, and high scores suggested a higher level of grateful disposition. Several researchers, including Froh et al. (2011), Solom et al. (2016), Toussaint and Friedman (2008), Van Dusen (2014), and Watkins et al. (2008), reported moderate to strong convergent and divergent validity of the scale.

A man high on lack of sense of deprivation or abundance appreciates his/her physical, psychological, and spiritual possessions (Garg et al., 2020). Such people do not feel cheated in life; instead, they enjoy abundance. A sense of abundance helps one to realize the fullness and richness of life instead of regretting the failure (Fredrickson, 2004). A simple appreciation trait is expressed as thankfulness and appreciation for routine and mundane events like meeting an old friend, awesome weather, and tasty food. An individual with a higher level of simple appreciation trait becomes thankful for most common aspects of life like a healthy human body, sound mind, steady income, good friends and relatives, etc. Further, a disposition of appreciation of others makes one acknowledge other's valuable contributions in one's life (Watkins et al., 2003).

GRAT-16 has been used extensively by previous researchers in both western and non-western contexts. In non-western settings, scholars like Naito and Sakata (2010), Chan (2010), and Garg et al. (2021) have relied on GRAT-16 to explore interrelationships between gratitude and other variables like happiness, vitality, and satisfaction with life. These studies were conducted in India (Garg, 2020; Mahipalan & Sheena, 2019), China (Kong et al., 2014), Japan (Naito & Sakata, 2010), and Vietnam (Tran et al., 2021). Similarly, western authors like Froh et al. (2011), Jans-Beken et al. (2015), Duran (2017), Hammer and Brenner (2017) have also used GRAT-16 for their empirical studies. Also, few researchers like Jans-Beken et al. (2015) and Duran (2017) have evaluated the adaptability of GRAT-16 in different cultural settings. Jans-Beken et al. (2015) reported acceptable internal consistency and test-rest reliability estimates among Dutch respondents. Also, good scores of Variance Inflation Index (VIF) and tolerance indicated no concern of multicollinearity. Duran (2017) concluded convergent and criterion validity and internal reliability of GRAT-16 among Turkish students. GRAT-16 score, LOSD, SA, and AO subscale scores were positively correlated with the Satisfaction with Life Scale (SWLS) with r values of 0.53, 0.52, 0.30, and 0.24, respectively (Duran, 2017). It was also reported that the GRAT-16 score, LOSD, SA, and AO subscale scores were negatively correlated with the Negative Affect Schedule score of Positive Affect and Negative Affect Scale (PANAS) with r-values of −0.25, −0.36, −0.08, and − 0.02, respectively. Similarly, Hammer and Brenner (2017) concluded acceptable psychometric properties of the scale in the American sample.

The inherent cultural differences in the conceptualization of gratitude (Naito & Washizu, 2015) require a psychometric analysis of a GRAT-16 scale in different cultures. A few studies like Mahipalan and Sheena (2019) and Garg (2020) have used GRAT-16 for Indian respondents. Mahipalan and Sheena (2019) reported a significant moderating effect of gratitude amid relationship of workplace spirituality and subjective happiness. Garg (2020) concluded sufficiency and necessity of gratitude for developing teacher leadership. Although these studies investigated the reliability and validity of the GRAT-16 scale, a detailed analysis of the scale's psychometric properties is required. Accordingly, the present study explores the reliability and validity of GRAT-16 in a sample of Indian college students. The three-pronged conceptualization of gratitude (in GRAT-16) seems well-grounded in the notion of *Ṛṇa* or pious obligations explained in Indian spiritual texts and scriptures. Indian spiritual *gurus* (teachers) considered the observance of 'pious obligations' called *Rnas* essential for spiritual upliftment and growth (Pandey & Navare, 2018). These sacred obligations (duties) are *Manuṣya-yajña* (towards fellow human beings), *Deva-yajña* (towards God and nature), *Pitr-Yajna (*towards the family, ancestors, and cultural values), *Brahma-yajña* (towards knowledge, skills, and talents), and *Bhūta-yajña* (towards the Eco-system). *Manuṣya-yajña* is the attitude of gratefulness for others for their contribution to one's life. This manifestation of thankfulness explains the 'appreciation for others' subscale of the GRAT-16. *Pitr-Yajna* is the expression of gratitude for ancestors, family, and friends. And *Brahma-yajña* encompasses thankfulness for acquired knowledge and skills, both in the realm of matter and spirit (Pandey & Navare, 2018). Family (genetics) and talent are two critical foundations of one's present material and spiritual possession. Thus, *Pitr-Yajna and Brahma-yajña* are an expression of gratitude for all assets of one's life. These two pious obligations seem to be covered under the 'sense of abundance/lack of sense of deprivation' subscale of SGRAT. *Bhūta-yajña* talks about duty towards nonhuman life, like plants, animals, birds, etc. And *Deva-yajña* is being grateful to God and nature for their blessings to humankind. It encourages prayers and spiritual ceremonies in honour of food, sun, water, wind, earth,

clouds, and mountains for their invaluable gifts to man. The 'simple appreciation' subscale of GRAT-16 seems suitably described by these two holy duties discussed in ancient Indian texts.

The scale may better suit Indian college students because the concept of 'pious obligations' is taught to students through both formal and informal systems. Courses on Business ethics, Technology ethics, Ethics and corporate social responsibility, etc., provide a theoretical and practical understanding of five *Rnas.* Informally, Indian parents, family members, and friends teach children to remain grateful in their lives through folk stories, dramas, and ancient rituals like daily prayers, touching the feet of parents, teachers, and elders, temple visits, etc. Also, India is a land of festivals, where each festival expresses thankfulness for one of the aspects of this beautiful life. These festivals inspire the younger generation to remain grateful in their lives. Thus, GRAT-16 seems to be notionally grounded in the Indian context also. It could be hypothesized that GRAT-16 is a valid measure of gratitude for the Indian students. The adaptability of GRAT-16 in the Indian context is investigated with the help of two studies. In the first study, exploratory and confirmatory factor analysis was conducted. And, in the second study, the reliability and validity of the scale were accessed.

Table 1 Mean descriptives of gratitude in sample 1 and sample 2 according to gender, age, qualification, university type, and course

Variable	Category	Sample 1		Sample 2	
		Mean	S.D	Mean	S.D
Gender	Male	6.36	1.414	6.46	1.585
	Female	6.54	1.357	6.51	1.432
Age	17–19 years	6.46	1.281	6.41	1.684
	20–22 years	6.32	1.354	6.48	1.690
	23–25 years	6.69	1.689	6.51	1.324
	Above 25 years	6.42	1.581	6.40	1.898
Qualification	UG	6.40	1.980	6.48	1.764
	PG	6.47	1.823	6.43	1.251
	PhD	6.60	1.900	6.52	1.746
	Other courses	6.64	1.642	6.50	1.649
University	Central	6.56	1.707	6.54	1.530
	State	6.53	1.402	6.44	1.932
	Private	6.16	1.958	6.39	1.271
Course	Commerce	6.43	1.315	6.42	1.840
	Science	6.52	1.883	6.48	1.695
	Humanities	6.45	1.495	6.70	1.849
	OVERALL	**6.45**	**1.648**	**6.47**	**1.853**

Source:- Primary Data

STUDY 1: Factorial Validity (Exploratory and Confirmatory Factor Analysis)

The first study examined the factorial validity of the three-dimensional scale of gratitude (GRAT-16) in the Indian context. It explored the construct validity of the GRAT-16 in the Indian context using an EFA and a CFA performed on two independent samples of students.

Method

Participants

A total of 530 students from different academic backgrounds participated in the present study. Participants from almost every state of India took part in the survey, with a majority from Delhi (46.6%). The sample comprised 51.32% male and 48.68% female students. The mean age of the respondents was 21.19 years. These students were enrolled in commerce (68.11%), science (20.57%), and humanities (11.32%) courses. Also, respondents registered in undergraduate, postgraduate, Ph.D., and other courses were 57.55%, 31.51%, 2.26%, and 8.68% of the sample. It was also observed that 23.77% of the participants were from the central universities, 53.21% from state universities, and 23.02% from private universities. The mean descriptive of the sample is shown in Table 1. The data was collected using a convenience sample. The participants were invited through online platforms to complete questionnaires created using Google forms.

Instruments

First, participants were enquired about their demographic characteristics (gender, age, residence state, qualification, type of courses, and universities). Second, Thomas and Watkins' (2003) Gratitude Resentment and Appreciation Scale (GRAT-16) was used to measure respondents' gratitude disposition. As discussed, the scale has three subscales: lack of a sense of deprivation or sense of abundance (six items), simple appreciation (six items), and appreciation for others (four items). The scale comprises sixteen statements. It is a nine-point rating scale ranging from 1 (strongly disagree) to 9 (strongly agree). Examples of statements are "oftentimes I have been overwhelmed at the beauty of nature," "life has been good to me." English version of GRAT-16 was used, as the Indian education system promotes multilingualism. The English language is introduced to the students from the very beginning of their academic life. Consequently, Indian students acquire an adequate understanding of the language till they join college to pursue their graduation and higher degrees. Thus, the original English version of the questionnaire was used in this study.

Procedure

First, the purpose of the study was clarified to the respondents. Second, they were ensured of confidentiality and anonymity of the data. They were promised that collected data would not be shared with anyone. Lastly, students were requested to read the statements and provide their honest feelings carefully. As a token of appreciation, all respondents were invited to join an online seminar on 'Art and science of gratitude.'

Data Analysis

The sample of 530 was divided into two parts with the help of systematic sampling. Even responses were added in the first subsample, and odd responses were included in the second subsample. However, it was observed that the second subsample had fewer respondents with a Ph.D. Thus, five respondents with the Ph.D. course were added in the second subsample. The first subsample (n1 = 260) was subjected to Exploratory Factor Analysis (EFA) using principal component analysis in SPSS statistics 21. It consisted of 50.76% male and 49.34% female students. Respondents enrolled in commerce, science, and humanities courses were 76.8%, 11.6%, and 11.6% of the sample. Also, 50.38% of students were pursuing undergraduate studies. Participants registered in postgraduate, doctorate, and other courses were 39.61%, 2.30%, and 7.71%, respectively. Before EFA, the Kaiser-Meyer-Olkin (KMO) test and Bartlett's test of sphericity were conducted. These tests define the suitability of data for factor analysis (factorability). The KMO measure of sampling adequacy was 0.840. Bartlett's test of sphericity was also significant at a 0.05 level of significance. These results concluded that the dataset could be subjected to EFA. Three factors having eigenvalue greater than one were extracted. Table 2 reports factor loadings, communalities, eigenvalues, and total variance explained by sixteen statements. Factor loadings ranged from 0.564 (statement- 07: "every fall, I really enjoy watching the leaves change colours") to 0.817 (statement-10: "more bad things have happened to me in my life than I deserve"). The results indicated acceptable factor loadings (more significant than |0.40|) for all sixteen statements. Sixteen statements explained 54.415% of the variance. The component matrix is depicted as annexure 1.

The second subsample (n2 = 270) was subjected to confirmatory factor analysis (CFA) using AMOS 18. This subsample comprised 51.85% male and 48.15% female students. Participants enrolled in undergraduate, postgraduate, doctorate, and other courses were 64.44%, 23.70%, 2.22%, and 9.62%, respectively. Also, 59.62% and 29.25% of respondents were enrolled in commerce and science courses. According to the literature, GFI ≥ 0.95; CFI ≥ 0.97; SRMR ≤0.05; and RMSEA <0.05 indicates a perfect fit. And, 0.90 ≤ GFI ≤ 0.95; 0.90 ≤ CFI ≤ 0.97; 0.05 ≤ SRMR ≤0.1, and 0.05 < RMSEA <0.08 reports acceptable fit (Kline, 2005; Şimşek, 2007). The default model showed some inconsistencies with poor values of RMSEA (0.091) and CFI (0.85). It highlighted the need for model revision. It was observed that the regression weight of one of the statements of LOSD (life has been good to me) was too low. The regression weight of the statement was 0.21. Thus, this statement was deleted from the model. Model fit indices were acceptable for the remaining fifteen statements ($\chi 2/df = 1.847$; GFI- 0.952, CFI = 0.914, SRMR = 0.038; RMSEA = 0.064). Figure 1 shows the revised and final model of GRAT-16. It concluded that although the finding supported a three-dimensional scale of GRAT-16, the deletion of one statement reduced the scale to a fifteen-item questionnaire. Subjected to reliability and validity testing, GRAT-15 may be proposed for the Indian population instead of GRAT-16. It may be referred to as the Indian GRAT-15 scale.

Study 2: Reliability and Validity Testing

Study 2 examined reliability using Cronbach's alpha and composite reliability values. It also investigated the convergent, discriminant, and construct validity of the GRAT-15.

Method

Participants

Study 2 was based on a sample of 519 students. There were 288 (53.56%) male participants and 241 (53.56%) female respondents. Also, 63.2%, 26.37%, and 10.31% of the participants belonged to commerce, science, and humanities streams. The mean value of the sample was 21.56 years. The students registered in central, state, and private universities were 23.77%, 53.21%, and 23.02% of the sample. Again, convenience sampling was used. A structured questionnaire was distributed among respondents through online platforms.

Instruments

First, demographic characteristics (gender, age, residence state, qualification, type of courses, and universities) of the participants were recorded. Then, fifteen statements, as derived after EFA and CFA, were used to measure gratitude. Diener et al. (2013). Satisfaction with Life Scale (SWLS) was used to measure participant's overall psychological satisfaction with life. It is a five-item scale that assessed psychological satisfaction on a seven-point scale, ranging from 1 (strongly disagree) to 7 (strongly agree). Few items of the scale are "in most ways my life is close to my ideal" and "the conditions of my life are excellent." Cronbach's alpha value of the scale for the present study was 0.844. Also, Kessler Psychological Distress Scale (K10), developed by

Table 2 Factor structure of GRAT-16

Items	Loadings	h^2	Eigen Value
Lack of Sense of Deprivation			**3.490**
"10 More bad things have happened to me in my life than I deserve. (R)	.817	.667	
03 There never seems to be enough to go around, and I never seem to get my share. (R)	.778	.606	
15 For some reason, I don't seem to get the advantages that others get. (R)	.770	.593	
02 Life has been good to me.	.742	.551	
11 Because of what I've gone through in my life, I really feel like the world owes me something. (R)	.738	.544	
06 I really don't think that I've gotten all the good things that I deserve in life. (R)	.727	.529	
Simple Appreciation			**2.980**
13 I think it's important to enjoy the simple things in life.	.781	.609	
16. I think it's important to appreciate each day that you are alive.	.757	.573	
09 I think that it's important to "Stop and smell the roses."	.744	.554	
12 I think that it's important to pause often to "count my blessings."	.713	.509	
04 Oftentimes, I have been overwhelmed at the beauty of nature.	.646	.417	
07 Every Fall I really enjoy watching the leaves change colors.	.564	.318	
Appreciation for Others			**2.236**
14 I feel deeply appreciative for the things others have done for me in my life.	.767	.589	
05 Although I think it's important to feel good about your accomplishments, I think that it's also important to remember how others have contributed to my accomplishments	.758	.574	
01 I couldn't have gotten where I am today without the help of many people.	.743	.552	
08 Although I'm basically in control of my life, I can't help but think about all those who have supported me and helped me along the way."	.722	.521	
Number of items	16		
Explained variance (%)	54.415		
Eigenvalues	8.707		

Source- Primary Data,* satisfactory factor loadings (i.e., greater than |.40|.), (*R*) reversed scored items, h^2 communality

Kessler et al. (2003), was used to determine the distress level of the respondents. It is a ten-item scale graded on a five-point scale ranging from 1 (none of the time) and 5 (all of the time). The overall score varies from 5 to 50. A higher score indicates higher levels of life psychological distress, and a lower score indicates a lower level of psychological distress. A few items of the scale are "in the past four weeks, about how often did you feel nervous?" and "in the past four weeks, about how often did you feel so nervous that nothing could calm you down?". Cronbach's Alpha for K10 was 0.902, which

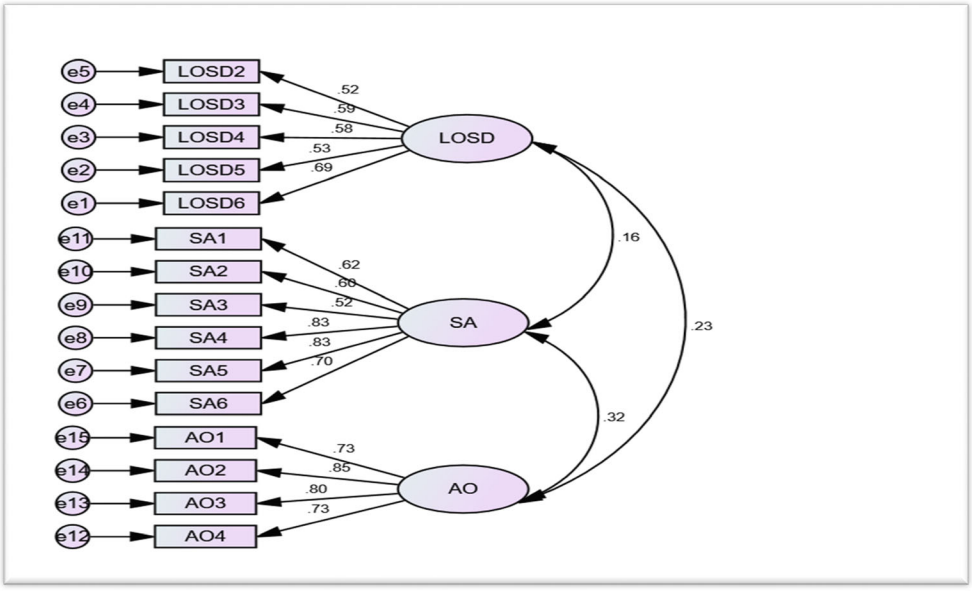


Fig. 1 Final Model of Indian GRAT-15. *Source: Primary Data*

confirmed the internal consistency reliability of the scale for the present study.

Procedure

Firstly, the participants were made aware of the purposes and objectives of the study. Second, they were assured of confidentiality and anonymity of the collected data. The respondents were promised that collected data would not be shared with a third party. The reliability of the Indian GRAT-15 scale was assessed with the help of Cronbach's alpha and composite reliability estimates. These values were calculated using SPSS 21. Although CFA ensured convergent and discriminant validity, they were again assessed using Average Variance Explained (AVE) values and the Fornell-Larcker testing method. Lastly, correlation coefficients between GRAT-15, SWLS, and K10 were calculated to verify the criterion validity of Indian GRAT-15.

According to George and Mallery (2003), Cronbach's alpha values should be equal to or greater than 0.70 to establish the internal consistency reliability of the scale. Cronbach's alpha is the average of all possible split-half correlations for a set of items. Later on, Cronbach and Shavelson (2004) suggested composite reliability values to re-establish the claim of reliability. Again, the acceptable value of the composite reliability estimate is 0.70. It was calculated using the below-mentioned formula-1. Convergent validity was investigated using average variance extracted (AVE), which was calculated with the help of the below-mentioned formula-2. According to Fornell and Larcker (1981), AVE values greater than 0.50 confirm the convergent validity of the scale.

$$\mathbf{CR}=(\sum \boldsymbol{\lambda i})^2 / \left((\sum \boldsymbol{\lambda i})^2+\sum\left(\mathbf{1}-\boldsymbol{\lambda}^2 \mathbf{i}\right)\right) \quad \text{—— Formula-1}$$

$$\mathbf{AVE}=\left(\sum \boldsymbol{\lambda}^2 \mathbf{i}\right)/\mathbf{N} \quad \text{—— Formula-2}$$

where is standardized factor loading for the item, N is the number of items in a factor.

Cronbach's alpha for Indian GRAT-15, LOSD, SA, and AO were 0.755, 0.764, 0.763, and 0.729, respectively. Also, composite reliability estimates were 0.947, 0.876, 0.854, and 0.845 for Indian GRAT-15, LOSD, SA, and AO. These values confirmed the good reliability of the Indian version of the scale. Also, acceptable values of AVE values were reported. The AVE values for Indian GRAT-15, lack of sense of deprivation, simple appreciation, and appreciation of others were 0.548, 0.587, 0.596, and 0.577. Table 3 shows the results of the Fornell and Larcker testing system for discriminant validity. It concluded that each construct's AVE values were greater than the squared correlation involving the constructs, which confirmed the discriminant validity of the scale. To test the criteria validity of the Indian GRAT-15, correlations coefficient between the Indian GRAT-15, SWLS and K10 were calculated. Correlation Coefficients between Gratitude, LOSD, SA, AO, SWLS, and K10 were depicted in Table 4**.** A positive correlation (r = 0.340, $p < .01$) was reported between the Indian GRAT-15 scores and SWLS scores. Also, correlation coefficients between LOSD, SA, and AO (subscale scores of the Indian GRAT-15) and the SWLS scores were 0.321, 0.320, and 0.321, respectively, at $p < .01$. It demonstrated that the participants with high levels of life satisfaction have high levels of gratitude. Also, the Indian GRAT-15 scores were negatively correlated with K10 scores (r = −0.359, p < .01). The LOSD, SA, and AO subscale scores were also negatively linked to K10 scores, with r values of −0.300, −0.192, −0.221 at p < .01. Thus, the Indian GRAT-15 was positively associated with SWLS and negatively linked with K10. These results concluded the criterion validity of the GRAT-15 in the Indian context.

Table 3 Fornell-Larcker Test for Discriminant Validity

Variables	LOSD	SA	AO
Lack of sense of deprivation	**1 (0.587)**		
Simple Appreciation	**0.09**	**1(0.596)**	
Appreciation for Others	**0.016**	**0.360**	**1(0.577)**

Source: Primary Data, AVE values are in bold (diagonally)

Discussion

Gratitude has a particular reference in Indian culture. Indian traditions and value systems represent the pinnacle of gratitude experience and manifestation (Garg, 2017). Indian epics, spiritual texts, customs, and practices recommend gratitude as a divine virtue. As an expression of gratefulness, Indians worship and honour almost everything for their valuable contribution to human life (Garg, 2018). Trees, mountains, birds, animals, rivers, earth, air, water, etc., are honoured for their usefulness. One of the most famous *mantras*, "*Om Namah*

Table 4 Correlation coefficients among gratitude, LOSD, SA, AO, SWLS, and K10 in a sample of 519 students of Study 2

Variables	GRAT-15	LOSD	SA	AO	SWLS	K10
GRAT-15	1					
LOSD	0.660	1				
SA	0.703	0.300	1			
AO	0.685	0.402	0.601	1		
SWLS	0.340	0.321	0.320	0.321	1	
K10	−0.359	−0.300	−0.192	−0.221	−0.262	1

Source: Primary Data

Shivaay," is an expression of thankfulness for God and others. Similarly, *Rig Veda's* verse, "Namo Bharanta Emasi" (verse 1.1.7), also shows a sense of honour and gratefulness for respectable. Thankfulness to God, nature, and fellow beings is an integral part of morning and evening prayer.

Although gratitude is everywhere in Indian epics and customs, scientific studies on the beneficial effects of gratitude are inadequate in Indian settings. The present study intends to initiate empirical exploration of gratitude among Indians through examining the generalizability of GRAT-16 for Indian respondents. For psychometric analysis of GRAT-16, two studies were conducted. Study 1 had two sub-samples. The first subsample was used for exploratory factor analysis using varimax rotation. The factorial solution reported a three-factor solution that explained 54% variance. The second subsample was used for confirmatory factor analysis to substantiate a three-factor scale. Here, one statement was deleted for low regression weight. It reduced GRAT-16 to GRAT-15 for Indian respondents with appreciable goodness of fit indices. The deleted item was 'life has been good for me.' Froh et al. (2011) also reported the removal of the same item owing to poor loading. Their findings were also based on a sample of students. The authors reasoned "vagueness of the statement" as a possible cause of its removal from the scale. It seems that Indian students also find it difficult to comprehend this statement. It means that the item is too general to elicit gratitude. Study 2 concluded good internal consistency reliability with acceptable values of Cronbach's alpha and composite reliability. The appropriate values of average variance explained confirmed convergent and discriminant validity of Indian GRAT-15. The criterion validity of the Indian GRAT-15 was also accessed. A significant positive correlation with the Satisfaction with Life Scale and a negative association with Kessler Psychological Distress Scale concluded that GRAT-15 might be a valuable tool to measure gratitude for Indian respondents.

The current study contributes to the existing literature and scientific knowledge of gratitude, often criticized for being dominated by western scholars. The notion of pious obligations (rnas) is still relevant in India (Pandey & Navare, 2018). The adaptability of GRAT-16 empirically substantiates the findings of the authors. As discussed, GRAT-16 matches the Indian notion of *rnas.* The generalization of the gratitude scale in the Indian context will motivate more empirical exploration on conceptualization, utilities, and limitations of gratitude. Psychologists, Human Resource Managers, and organization development consultants may use GRAT-15 to measure gratitude among Indian respondents. The findings also suggest that gratitude correlates positively with satisfaction with life. It also reports a statistically significant negative correlation between gratitude and stress level. It concludes that the experience of gratitude may increase one's satisfaction level and decrease stress level. Although the study offers exciting findings, it has several limitations too. Firstly, the results are based on a sample of students only; thus, findings may not be generalized to other age groups. Future researchers may use an age-diversified sample set for the generalization of results for the entire Indian population. Secondly, only internal consistency reliability is measured in this study. Future studies may also explore test-retest or interrater reliability. Future researchers may also examine external and content validity. Also, a cross-cultural survey of gratitude may compare gratefulness among respondents of different cultural backgrounds.

Annexure

Table 5 Component matrix

Statements	1	2	3
Statement 1	.817	.274	−.294
Statement 2	.778	.077	.224
Statement 3	.770	.132	−.095
Statement 4	.742	.295	.223
Statement 5	.738	.213	.125
Statement 6	.727	−.093	.125
Statement 7	.324	.781	.212
Statement 8	.093	.757	.165
Statement 9	−.222	.744	.158
Statement 10	.147	.713	.285
Statement 11	.390	.646	.371
Statement 12	.047	.564	.158
Statement 13	.114	.295	.767
Statement 14	−.084	.395	.758
Statement 15	.312	.320	.743
Statement 16	.211	−.075	.722

Source:-Primary Data

Declaration

a) there is no conflict of interest
b) No funding from any source
c) Respondents were duly informed of the study's purpose, and consent was procured from all the respondents.

Also, we declare that the datasets generated during and/or analyzed during the current study are not publicly available due promise of confidentiality to respondents.

References

Algoe, S. B. (2012). Find, remind, and bind: The functions of gratitude in everyday relationships. *Social and Personality Psychology Compass, 6*(6), 455–469. https://doi.org/10.1111/j.1751-9004.2012.00439.x

Algoe, S. B., & Zhaoyang, R. (2015). Positive psychology in context: Effects of expressing gratitude in ongoing relationships depend on perceptions of enactor responsiveness. *The Journal of Positive Psychology, 11*(4), 399–415. https://doi.org/10.1080/17439760.2015.1117131

Chan, D. W. (2010). Gratitude, gratitude intervention and subjective well-being among Chinese school teachers in Hong Kong. *Educational Psychology, 30*(2), 139–153. https://doi.org/10.1080/01443410903493934

Chan, D. W. (2011). Burnout and life satisfaction: Does gratitude intervention make a difference among Chinese school teachers in Hong Kong? *Educational Psychology, 31*(7), 809–823. https://doi.org/10.1080/01443410.2011.608525

Cronbach, L. J., & Shavelson, R. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement, 64*(3), 391–418. https://doi.org/10.1177/0013164404266386

Diener, E., Inglehart, R., & Tay, L. (2013). Theory and validity of life satisfaction measures. *Social Indicators Research, 112*(3), 497–527. https://doi.org/10.1007/s11205-012-0076-y

Drążkowski, D., Kaczmarek, L. D., & Kashdan, T. B. (2017). Gratitude pays: A weekly gratitude intervention influences monetary decisions, physiological responses, and emotional experiences during a trust-related social interaction. *Personality and Individual Differences, 110*, 148–153. https://doi.org/10.1016/j.paid.2017.01.043

Duran, N. O. (2017). The revised short gratitude, resentment, and appreciation test (S-GRAT): Adaptation for Turkish college students. *Journal of Happiness & Well-being, 5*(1), 23–37.

Emmons, R. A., & McCullough, M. E. (2003). Counting blessings versus burdens: An experimental investigation of gratitude and subjective well-being in daily life. *Journal of Personality and Social Psychology, 84*(2), 377–389. https://doi.org/10.1037/0022-3514.84.2.377

Emmons, R. A., & Mishra, A. (2011). Why gratitude enhances well-being: What we know, what we need to know. In K. M. Sheldon, T. B. Kashdan, & M. F. Steger (Eds.), *Designing positive psychology: Taking stock and moving forward (pp. 248–262)*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195373585.003.0016

Fehr, R., Fulmer, A., Awtrey, E., & Miller, J. A. (2017). The grateful workplace: A multilevel model of gratitude in organizations. *Academy of Management Review, 42*(2), 361–381. https://doi.org/10.5465/amr.2014.0374

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.1177/002224378101800104

Fredrickson, B. L. (2004). Gratitude, like other positive emotions, broadens and builds. In R. A. Emmons & M. E. McCullough (Eds.), *The psychology of gratitude (pp. 145–166)*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195150100.003.0008

Froh, J. J., Fan, J., Emmons, R. A., Bono, G., Huebner, E. S., & Watkins, P. (2011). Measuring gratitude in youth: Assessing the psychometric properties of adult gratitude scales in children and adolescents. *Psychological Assessment, 23*(2), 311–324. https://doi.org/10.1037/a0021590

Garg, N. (2017). Workplace Spirituality and Organizational Commitment: An Empirical Exploration. *Purushartha - A Journal of Management , Ethics and Spirituality*, *10*(02). https://doi.org/10.21844/pajmes.v10i02.10566

Garg, N. (2020). Exploring role of gratitude in developing teacher leadership in Indian universities. *International Journal of Educational Management, 34*(5), 881–901. https://doi.org/10.1108/ijem-07-2019-0253

Garg, N., & Sarkar, A. (2020). Vitality among university students: Exploring the role of gratitude and resilience. *Journal of Organizational Effectiveness: People and Performance, 7*(3), 321–337. https://doi.org/10.1108/joepp-03-2020-0041

Garg. (2018). Promoting organizational performance in Indian insurance industry: The roles of workplace spirituality and organizational citizenship behaviour. *Global Business Review, 21*(3), 834–849. https://doi.org/10.1177/0972150918778983

Garg, N., Gera, S., & Punia, B. K. (2021). Gratitude and turnover intention in E-commerce startups. *Journal of Electronic Commerce in Organizations, 19*(2), 34–54. https://doi.org/10.4018/jeco.2021040103

Garg, N., Jain, A., & Punia, B. (2020). Gratitude, social intelligence, and leadership among university teachers: Mediation and moderation analysis. *International Journal of Organizational Analysis, 29*(2), 368–388. https://doi.org/10.1108/ijoa-01-2020-2016

Garg, N., Punia, B. K., & Jain, A. (2019). Workplace spirituality and job satisfaction: Exploring mediating effect of organization citizenship behaviour. *Vision: The Journal of Business Perspective, 23*(3), 287–296. https://doi.org/10.1177/0972262919850928

George, D., & Mallery, P. (2003). *SPSS for windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Allyn & Bacon.

Grant, A. M., & Gino, F. (2010). A little thanks goes a long way: Explaining why gratitude expressions motivate prosocial behavior. *Journal of Personality and Social Psychology, 98*(6), 946–955. https://doi.org/10.1037/a0017935

Grant, A. M., & Wrzesniewski, A. (2010). I won't let you down. . . or will I? Core self-evaluations, other-orientation, anticipated guilt and gratitude, and job performance. *Journal of Applied Psychology, 95*(1), 108–121. https://doi.org/10.1037/a0017974

Hammer, J. H., & Brenner, R. E. (2017). Disentangling gratitude: A theoretical and psychometric examination of the gratitude resentment and appreciation test–revised short (GRAT–RS). *Journal of Personality Assessment, 101*(1), 96–105. https://doi.org/10.1080/00223891.2017.1344986

Hill, P. L., & Allemand, M. (2011). Gratitude, forgivingness, and well-being in adulthood: Tests of moderation and incremental prediction. *The Journal of Positive Psychology, 6*(5), 397–407. https://doi.org/10.1080/17439760.2011.602099

Jans-Beken, L., Lataster, J., Leontjevas, R., & Jacobs, N. (2015). Measuring gratitude: A comparative validation of the Dutch gratitude questionnaire (GQ6) and short gratitude, resentment, and appreciation test (SGRAT). *Psychologica Belgica, 55*(1), 19–31. https://doi.org/10.5334/pb.bd

Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S. L. T., Manderscheid, R. W., Walters, E. E., & Zaslavsky, A. M. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry, 60*(2), 184–189. https://doi.org/10.1001/archpsyc.60.2.184

Kline, R. B. (2005). *Principles and practice o f structural equation modeling* (2nd ed.). Guilford Press.

Kong, F., Ding, K., & Zhao, J. (2014). The relationships among gratitude, self-esteem, social support and life satisfaction among undergraduate students. *Journal of Happiness Studies, 16*(2), 477–489. https://doi.org/10.1007/s10902-014-9519-2

Krause, N., & Hayward, R. D. (2014). Hostility, religious involvement, gratitude, and self-rated health in late life. *Research on Aging, 36*(6), 731–752. https://doi.org/10.1177/0164027513519113

Mahipalan, M., & Sheena, S. (2019). Workplace spirituality and subjective happiness among high school teachers: Gratitude as a moderator. *Explore, 15*(2), 107–114. https://doi.org/10.1016/j.explore.2018.07.002

McCullough, M. E., Emmons, R. A., & Tsang, J. A. (2002). The grateful disposition: A conceptual and empirical topography. *Journal of*

Personality and Social Psychology, 82(1), 112–127. https://doi.org/10.1037/0022-3514.82.1.112

McCullough, M. E., Kilpatrick, S. D., Emmons, R. A., & Larson, D. B. (2001). Is gratitude a moral affect? *Psychological Bulletin, 127*(2), 249–266. https://doi.org/10.1037/0033-2909.127.2.249

Michie, S. (2009). Pride and gratitude. *Journal of Leadership & Organizational Studies, 15*(4), 393–403. https://doi.org/10.1177/1548051809333338

Mills, P. J., Redwine, L., Wilson, K., Pung, M. A., Chinh, K., Greenberg, B. H., Lunde, O., Maisel, A., Raisinghani, A., Wood, A., & Chopra, D. (2015). The role of gratitude in spiritual well-being in asymptomatic heart failure patients. *Spirituality in Clinical Practice, 2*(1), 5–17. https://doi.org/10.1037/scp0000050

Millstein, R. A., Celano, C. M., Beale, E. E., Beach, S. R., Suarez, L., Belcher, A. M., Januzzi, J. L., & Huffman, J. C. (2016). The effects of optimism and gratitude on adherence, functioning and mental health following an acute coronary syndrome. *General Hospital Psychiatry, 43*, 17–22. https://doi.org/10.1016/j.genhosppsych.2016.08.006

Morgan, B., Gulliford, L., & Kristjánsson, K. (2014). Gratitude in the UK: A new prototype analysis and a cross-cultural comparison. *The Journal of Positive Psychology, 9*(4), 281–294. https://doi.org/10.1080/17439760.2014.898321

Naito, T., & Sakata, Y. (2010). Gratitude, indebtedness, and regret on receiving a friend's favour in Japan. *Pyschologia, 53*(3), 179–194. https://doi.org/10.2117/psysoc.2010.179

Naito, T., & Washizu, N. (2015). Note on cultural universals and variations of gratitude from an east Asian point of view. *The Journal of Behavioral Science, 10*(2), 1–8. https://doi.org/10.14456/ijbs.2015.38

Pandey, A., & Navare, A. V. (2018). Paths of yoga: Perspective for workplace spirituality. The Palgrave Handbook of Workplace Spirituality and Fulfillment, 101–126.

Sarkar, A., & Garg, N. (2020). "Peaceful workplace" only a myth? Examining the mediating role of psychological capital on spirituality and nonviolence behaviour at the workplace. *International Journal of Conflict Management, 31*(5), 709–728. https://doi.org/10.1108/ijcma-11-2019-0217

Sarkar, A., & Garg, N. (2021). A multilevel model exploring the relationships between individual spirituality, spiritual climate, and nonviolent work behavior. *Conflict Resolution Quarterly, 38*(3), 209–229. https://doi.org/10.1002/crq.21298

Şimşek, Ö. F. (2007). *YapısalEşitlikModellemesinegiriş: Temel ilkelerveLisreluygulamaları*. EkinoksYayınları.

Solom, R., Watkins, P. C., McCurrach, D., & Scheibe, D. (2016). Thieves of thankfulness: Traits that inhibit gratitude. *The Journal of Positive Psychology, 12*(2), 120–129. https://doi.org/10.1080/17439760.2016.1163408

Sommers, S., & Kosmitzki, C. (1988). Emotion and social context: An American-German comparison. *British Journal of Social Psychology, 27*(1), 35–49. https://doi.org/10.1111/j.2044-8309.1988.tb00803.x

Thomas, N. & Watkins, P. (2003). Measuring the grateful trait: Development of the revised GRAT. Poster presented to the annual convention of the Western psychological association, may, 2003, Vancouver, BC.

Toussaint, L., & Friedman, P. (2008). Forgiveness, gratitude, and well-being: The mediating role of affect and beliefs. *Journal of Happiness Studies, 10*(6), 635–654. https://doi.org/10.1007/s10902-008-9111-8

Tran, T. A. T., Nguyen Phuoc, C. T., Dinh, H. V. T., & Nguyen, V. T. (2021). The Vietnamese version of the gratitude questionnaire (GQ) and the revised short gratitude, resentment, and appreciation test (S-GRAT): Psychometric properties among adolescents. *Japanese Psychological Research*. Published. https://doi.org/10.1111/jpr.12321.

Tudge, J. R., Freitas, L. B., & O'Brien, L. T. (2015). The virtue of gratitude: A developmental and cultural approach. *Human Development, 58*(4–5), 281–300. https://doi.org/10.1159/000444308

Van Dusen, J. P. (2014). Relationships amongst gratitude, well-being and depression (Doctoral dissertation, University of Toledo).

Watkins, P. C., Cruz, L., Holben, H., & Kolts, R. L. (2008). Taking care of business? Grateful processing of unpleasant memories. *The Journal of Positive Psychology, 3*(2), 87–99.

Watkins, P. C., Uhder, J., & Pichinevskiy, S. (2014). Grateful recounting enhances subjective well-being: The importance of grateful processing. *The Journal of Positive Psychology, 10*(2), 91–98. https://doi.org/10.1080/17439760.2014.927909

Watkins, P. C., Woodward, K., Stone, T., & Kolts, R. L. (2003). Gratitude and happiness: Development of a measure of gratitude and relationship with subjective well-being. *Social Behavior and Personality: An International Journal, 31*(5), 431–451. https://doi.org/10.2224/sbp.2003.31.5.431

Wood, A. M., Froh, J. J., & Geraghty, A. W. (2010). Gratitude and well-being: A review and theoretical integration. *Clinical Psychology Review, 30*(7), 890–905. https://doi.org/10.1016/j.cpr.2010.03.005

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

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

Implementation of High Performance 4-Bit ALU using Dual Mode Pass Transistor Logic

Anurag Chauhan
Department of Electronics and Communication Engineering
Delhi Technological University
(formerly Delhi College of Engineering)
Delhi, India
anuragchauhan@dtu.ac.in

Kirandeep Kaur Saini
Department of Electronics and Communication Engineering
Delhi Technological University
(formerly Delhi College of Engineering)
Delhi, India
kkirandeep190@gmail.com

Nitin Rajput
Department of Electronics and Communication Engineering
Delhi Technological University
(formerly Delhi College of Engineering)
Delhi, India
rajputnitin678@gmail.com

Rushil Domah
Department of Electronics and Communication Engineering
Delhi Technological University
(formerly Delhi College of Engineering)
Delhi, India
rushildomah@gmail.com

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498553

***Abstract*— In this paper, we present a four bit arithmetic logic unit which is energy efficient and temperature invariant implemented using the dual mode pass transistor logic. The basic logic gates such as NOR and NAND are designed using both CMOS logic and dual mode pass transistor logic and are used in the proposed design. Simulations performed demonstrated that DMPL can reduce the computed worst case delay by 42.39% and 39.13% for NOR and NAND gates respectively in dynamic mode and average power dissipation by 67.96% and 24.09% for NOR and NAND gates respectively in static mode. In the implemented Arithmetic and Logic Unit, we observe a reduction in worst case delay and average power dissipation by 62.67% and 28.28%. The proposed logic was implemented in 90nm bulk technology using Cadence® Virtuoso® Schematic Editor.**

Keywords*— *Dual Mode Pass Transistor Logic (DMPL), energy efficiency, power-delay product.

I. INTRODUCTION

An arithmetic logic unit (ALU) is a basic building block of a processing unit and is a combinational digital circuit which performs logical and arithmetic operations [9], on binary numbers. It is also termed as a vital part of most computing devices. The performance of an ALU is determined by metrics like delay and average power dissipation [2].

To implement this circuit, we use the Dual Mode Pass Transistor Logic (DMPL) because of its intrinsic advantages of providing high speed, low power dissipation and the ability to switch between different modes of operation [1]. To substantiate the performance of the design proposed, we implemented a four bit arithmetic and logic unit using the various building blocks in 90nm bulk technology and compared metrics such as delay and average power dissipation. The simulation results show that the proposed Arithmetic and Logic Unit architecture has a speed up of 62.67% in dynamic mode and reduction in average power dissipation (at 25ºC) by 28.28% in the static mode of operation.

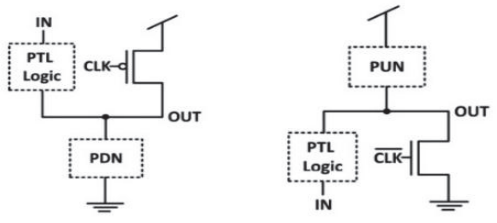
The remainder of this brief is constructed as follows: Section II introduces the proposed DMPL logic used and presents NOR and NAND implementation. Section III describes the Arithmetic Logic Unit (ALU) construction and the logics to be implemented. Simulation outputs of the different functionality blocks and schematic of ALU in both CMOS and DMPL are given in Section IV. Section V comprises results and graphs of average power dissipation resulting from temperature variation. Section VI concludes the brief.

II. DUAL MODE PASS LOGIC

Even if the dual mode logic (DML) allows for interchangeable operation of logic circuits between static and dynamic modes and provides low granularity power-delay optimization at gate levels [1], it exhibited a tradeoff between higher power consumption and higher speed.

DMPL, upon combining attributes of both pass transistor logic and dual mode logic allows low power dissipation in static mode and faster operation in the dynamic mode. The unique selling proposition of this logic shows power and speed improvement as compared to DML along with the ability to switch between the two modes of operation.

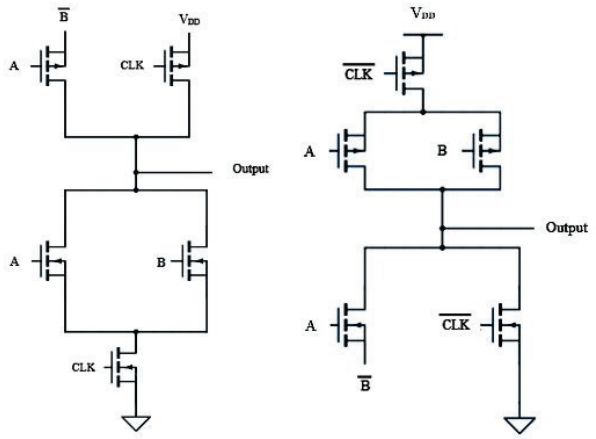
The structure of basic DMPL, both type A and type B [1] shown in Fig 1 is described as follows: it consists of a minimum sized PTL logic which substitutes the stacked up transistors in CMOS logic (pull up in Type-A and pull down in Type-B). In parallel to the PTL, a clocked transistor is added, this effectively reduces the intrinsic capacitance leading to speed up and increased energy efficiency.



(a) (b)

Fig. 1. . (a) Type A and (b) Type B topology of DMPL [1]

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:02:47 UTC from IEEE Xplore. Restrictions apply.



(a) (b)

Fig. 2. . (a) Type A NOR with footer and (b) Type B NAND with header gates in DMPL.

NAND and NOR gates implemented in DMPL are shown in Fig 2. Considering Type-B NAND , the series transistors in PDN are replaced by a minimum sized PTL logic, the gate of which is connected to input A and source of PTL transistor is connected to $\bar{B}$, in parallel with this PTL logic we connect a clocked transistor to switch between dynamic and static modes of operation.

In Static mode, the value to '*clk_bar'* input is stable at '0' whereas in Dynamic mode there is switching between '1' and '0' for pre-discharge and evaluate phase respectively. A header is also connected in the PUN to avoid any glitches in dynamic mode of operation [2].

Similar approach is adopted while implementing Type-A NOR with a footer transistor in PDN. DMPL is thus said to reduce the node capacitance by using minimum sized transistors.

III. Proposed Arithmetic and Logic Unit

The architecture of the proposed Arithmetic and Logic Unit given in Figure 3 is implemented mainly using headed and footed DMPL gates.

While design with the DMPL is very intuitive, the few design considerations employed in the design of the proposed ALU are as follows:

(i) There were certain limitations while cascading dynamic logics which are subdued using Domino logic and NP logic [6]. Using these two as inspiration, we came to our ALU implementation in which we are cascading Type A and Type B blocks.

(ii) While cascading using DMPL, we ensure that any input to Type-B block comes from a Type-A block and vice versa. This step consequently prevents short circuit currents, cascading issues and loss of charge [2].

(iii) It is ensured that all blocks of both Type-A and Type-B evaluate in the same phase of the clock [2].

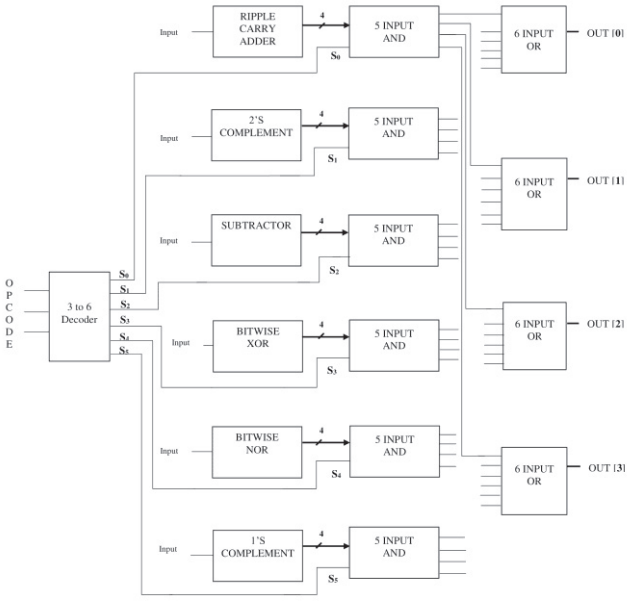


Fig. 3. Arithmetic and Logic Unit Block Diagram.

The implemented functionalities of the given Arithmetic and Logic Unit are depicted in Table 1.

TABLE I. ALU Functionality Table

Opcode	***Function, F***	***Functionality***
000	A + B	4 bit addition
001	$\bar{A}$ + 1	2's complement of A
010	A - B	4 bit subtraction
011	A ^ B	Bitwise XOR of A&B
100	(A + B)'	Bitwise NOR of A&B
101	$\bar{A}$	1's complement of A

IV. Setup and Simulations

All circuits implemented in static CMOS and DMPL were simulated in Cadence® Virtuoso® using the 90nm bulk technology. All simulations were carried out considering worst case delays and at V_{DD}=1.0V. The schematics of ALU implemented in CMOS logic and DMPL are given in Figure 4 and 6 respectively. Figures 5 and 7 illustrate the simulations of the implemented ALU using CMOS and DMPL.

The waveform shows the inputs A and B collapsed into a 4 bit bus and corresponding to the chosen opcode we obtain an output, where sum and difference are presented in signed decimal format and the rest are represented in binary form.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:02:47 UTC from IEEE Xplore. Restrictions apply.

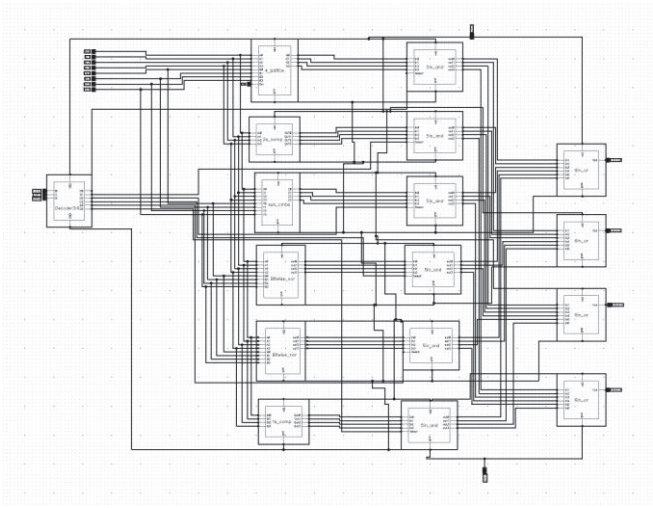


Fig. 4. ALU schematic implemented in CMOS



Fig. 5. Outputs of the various functionalities in CMOS implementation.

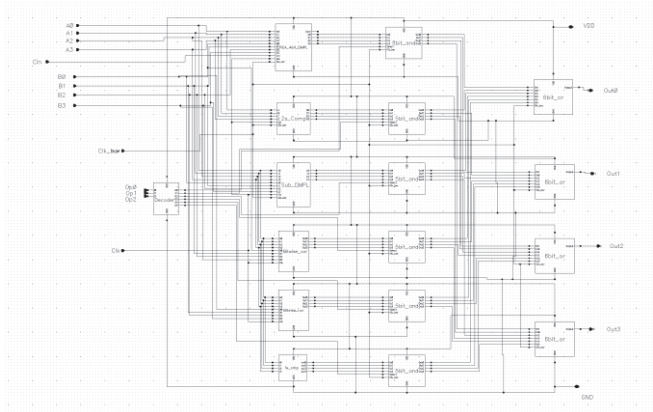


Fig. 6. ALU schematic implemented in DMPL.



Fig. 7. Outputs of functionality blocks implemented using DMPL.

V. Results and Interpretation

Tables 2-7 illustrate the results obtained. Table 2-4 shows the average propagation delay, average power dissipation and power delay product in NOR and NAND gates in CMOS, Dynamic DMPL and Static DMPL Implementations respectively. From Table. 2, the results obtained shows a speed up of 42.39% in NOR and 39.13% in NAND in DMPL dynamic mode. This improvement from CMOS to DMPL results from the fact that charging/discharging of node capacitances (CL) in CMOS takes place through a series of transistors which are larger in size whereas in DMPL, the same is happening through a single PTL transistor which is minimum sized. The delay is further reduced due to the reduction of node capacitances (CL) due to the use of minimum sized transistors.

The simulation results obtained for the gates when the variation in temperature is considered are given in Table 3. The outputs are recorded at 25ºC, 50ºC, 75ºC and 100ºC. Graphical representations of the same are provided in Figure 8 (a) and (b). Considering a temperature of 25ºC, the power savings obtained when NOR and NAND gates are operated in the static mode are 67.96% and 24.09% respectively. We observe a reduction in power dissipation because upon using minimum sized transistors, the C_L decreases which eventually decreases the switching power. A general observation from Table 3 is that as temperature increases, the percentage increase in power dissipation in DMPL is lesser than CMOS. Table 4 illustrates the power delay products of NOR and NAND gates respectively.

Table 5-7 depicts the average delay, average power dissipation at various temperatures and the power delay product for the implemented Arithmetic and Logic Unit (ALU). From Table 5 and 6, we see that the ALU has a speed up of 62.67% in dynamic mode and reduction in average power dissipation (at 25ºC) by 28.28% in the static mode of operation. The total transistor count for the gates and ALU is given in Table 8. Although more transistors are being used when ALU is implemented in DMPL as compared to CMOS, however the area overhead is not much because of minimum sized Pass Transistor Logic and clock transistor.

TABLE II. Average Propagation Delay in NOR and NAND gates

Gate Type	Average Propagation Delay (ps)		
	CMOS	***DMPL Dynamic***	***DMPL Static***
NOR	72.46	41.72	58.69
NAND	64.22	39.09	53.45

TABLE III. Average Power Dissipation in NOR and NAND gates

Gate type	Temperature (ºC)	Average Power Dissipation (µW)		
		CMOS	***DMPL Dynamic***	***DMPL Static***
NOR	25	1.765	2.220	0.5654
	50	4.665	4.388	0.7245
	75	7.4277	7.284	1.240
	100	12.1071	10.555	2.179
NAND	25	1.693	1.869	1.2866
	50	3.555	4.157	2.7018
	75	6.7545	7.1516	5.538
	100	11.4826	10.458	7.923

TABLE IV. Power Delay Product of NOR and NAND gates.

Gate type	Power Delay Product (fJ)		
	CMOS	***DMPL Dynamic***	***DMPL Static***
NOR	0.127	0.0926	0.03318
NAND	0.108	0.07306	0.06876

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:02:47 UTC from IEEE Xplore. Restrictions apply.

TABLE V. AVERAGE DELAY IN IMPLEMENTED ALU.

Cell	Average Delay (ps)		
	CMOS	***CMOS***	***CMOS***
ALU	458.7	458.7	458.7

TABLE VI. AVERAGE POWER DISSIPATION IN ALU FOR VARIOUS TEMPERATURES.

Cell	Temperature (°C)	Average Power Dissipation (µW)		
		CMOS	***DMPL Dynamic***	***DMPL Static***
ALU	25	19.8	51.69	14.2
	50	37.62	84.8254	27.974
	75	79.002	144.075	49.51
	100	181.704	257.895	67.33

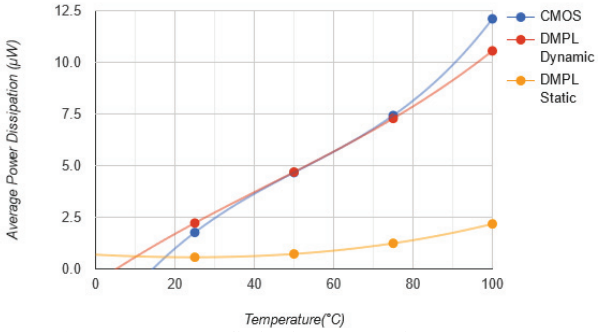
TABLE VII. POWER DELAY PRODUCT OF ALU AT 25°C AND VDD = 1.0V

Cell	Power Delay Product (fj)		
	CMOS	***DMPL Dynamic***	***DMPL Static***
ALU	9.082	8.849	5.912

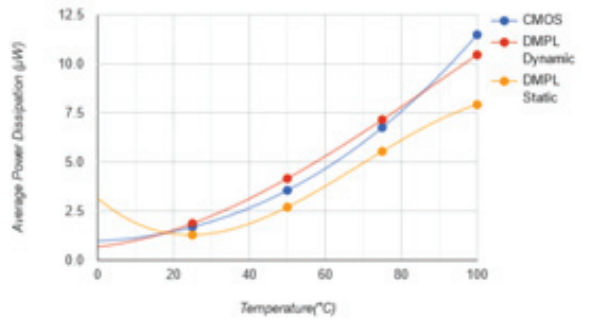
TABLE VIII. TRANSISTOR COUNT IN EACH IMPLEMENTATION.

Cell	Transistor Count	
	CMOS	***DMPL***
NOR & NAND	4	5
ALU	986	1242

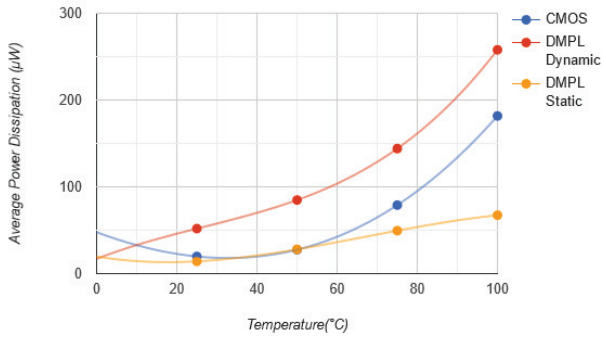
The graphs showing the resulting average power dissipations with varying temperature for NOR gate, NAND gate and the implemented Arithmetic Logic Unit are depicted in Figure. 8. (a), (b) and (c) respectively.



(a) Average power dissipation vs Temp in NOR gate



(b) Average power dissipation vs Temp in NAND gate



(c) Average power dissipation vs Temp in ALU

Fig. 8. Average power dissipation for varying temperature in (a) NOR gate (b) NAND gate and (c) ALU.

VI. CONCLUSION

Circuits designed using the Dual Mode Pass Transistor Logic (DMPL) exhibited smaller worst case propagation delays compared to those implemented in static CMOS logic. We observe significantly lower power delay product in DMPL than CMOS which is due to inherent lower propagation delay. For real time processing, the ALU implemented by DMPL can be tuned onto dynamic operation for commendable performance while those not requiring fast processing times can be switched to static operation. DMPL offers improvements in performance and energy savings despite using an increased number of transistors to implement the same functionalities as compared to other logics. The graphs give indication that DMPL is much immune to temperature variations as compared to CMOS when temperature is varied.

REFERENCES

[1] I. Stanger, N. Shavit, R. Taco, L. Yavits, M. Lanuzza and A. Fish, "Robust Dual Mode Pass Logic (DMPL) for Energy," 2020 IEEE International Symposium on Circuits and Systems (ISCAS), Seville, Spain , 2020.

[2] J. M. Rabaey, A. P. Chandrakasan and B. Nikolic, "Digital integrated circuits: a design perspective," Upper Saddle River, N.J.: Pearson Education, 2003.

[3] I. Levi, A. Kaizerman, and A. Fish, "Low voltage dual mode logic: Model analysis and parameter extraction," Microelectronics journal, vol.44, no. 6, pp. 553-560, 2013.

[4] N.Shavit, R.Taco, and A.Fish "Efficiency of dual mode logic in nanoscale technology nodes, " in 2018 IEEE International Conference on Science of Electrical Engineering in Israel(ICSEE). IEEE, 2018, pp. 1-4.

[5] A. Kaizerman, S. Fisher, and A. Fish, "Subthreshold dual mode logic," IEEE Transactions on Very Large Scale Integration(VLSI) Systems, vol. 21, no. 5, pp. 979–983, 2012.

[6] I.Levi, O.Bass, A. Kaizerman, A. Belenky and A.Fish, "High speed dual mode logic carry look ahead adder," in 2012 IEEE International Symposium on Circuits and Systems. IEEE, 2012, pp. 3037-3040.

[7] L. Mahor, A. Chauhan and P. Tiwari, "Parity Preserving Reversible Design Using FinFETs," 2019 1st International Conference on Signal Processing, VLSI and Communication Engineering (ICSPVCE), Delhi, India, 2019, pp. 1-6.

[8] M. Alioto, "Ultra-Low Power VLSI Circuit Design Demystified and Explained: A Tutorial," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 59, no. 1, pp. 3-29, Jan. 2012.

[9] J. L. Hennessey and D. A. Patterson, "Computer Architecture. A Quantitative Approach," San Francisco: Morgan Kaufmann, 1990.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:02:47 UTC from IEEE Xplore. Restrictions apply.

Title Page

Title:

Intelligent Algorithm based Efficient Planning of Electric Vehicle Charging Station: A Case Study of Metropolitan City of India

M. Bilal[1]*, Mohammad Rizwan[1,2]
[1]Department of Electrical Engineering, Delhi Technological University, Delhi-110042, India.
[2]Department of Electrical Engineering, College of Engineering, Qassim University, Buraydah 52571, Qassim, Saudi Arabia
Corresponding Author: *bilal.zhcet01@gmail.com, Mob No- +91 8279317577

Author names and affiliations:

1. Author **(Corresponding author)**: MOHD BILAL
   Designation: Ph.D. Scholar
   Department of Electrical Engineering
   Delhi Technological University
   Delhi - 110042, India
   Email: bilal.zhcet01@gmail.com
   ORCID: 0000-0003-1759-5138

2. Co-author: MOHAMMAD RIZWAN
   Designation: Professor
   Department of Electrical Engineering
   Delhi Technological University
   Delhi - 110042, India
   Email: rizwan@dce.ac.in
   ORCID: 0000-0003-2839-7241

Corresponding Author's Details:

MOHD BILAL
PhD Scholar
Department of Electrical Engineering
Delhi Technological University
Delhi - 110042, India
Mobile: +91 8279317577
Primary Email: bilal.zhcet01@gmail.com
Institute Email: bilal_2k17phdee05@dtu.ac.in

Intelligent Algorithm based Efficient Planning of Electric Vehicle Charging Station: A Case Study of Metropolitan City of India

M. Bilal[1]*, Mohammad Rizwan[1,2]
[1]Department of Electrical Engineering, Delhi Technological University, Delhi-110042, India.
[2]Department of Electrical Engineering, College of Engineering, Qassim University, Buraydah 52571, Qassim, Saudi Arabia
Corresponding Author: *bilal.zhcet01@gmail.com, Mob No- +91 8279317577

Abstract

Electric vehicles (EVs) are gaining popularity and becoming the mainstream mode of transportation in urban and rural areas in India and globally in the next few years. In the adoption of EVs, there are certain issues like proper charging infrastructure, charging time etc. and out of these the sizing and siting of the charging station (CS) particularly in urban areas where the cost of land and location plays an important role. Thus, it is important that the CS location be easily accessible for the EV users and cost-effective. This paper presents an intelligent algorithm based efficient planning of electric vehicle charging stations (EVCS) considering geographical information and road network. The cost function has been considered as investment cost, charging station electrification cost, electric vehicle energy loss cost, and travel time cost. An intelligent algorithm-based approach is employed to solve the planning problem of EVCS. Further, the impact on the reliability of the grid is also evaluated by determining the charging cost loss on each considered location. The result shows that the applied method provides better-optimized solutions which are beneficial for EV users, CS operators and the utility grid.

Keywords*:*

Electric Vehicles; Smart Transportation; Grey Wolf Optimization; Electric Vehicles Charging Infrastructure; Smart Grid.

1. Introduction

Nowadays, Electric vehicles (EVs) are gaining popularity because of reduction in consumption of fossil fuels and high demand in power sector [1, 2]. Cost-effective and eco-friendly natures of EVs are the major concerns which lead to deployment of EV on a large-scale [3, 4]. EVs can be categorized into battery EV and hybrid EV [5, 6]. Electric vehicle charging stations (EVCS) are installed to supply electricity for charging EVs. Capacity of EV battery ranges in between 20 to 60 kWh. Number of charging levels are available for EV charging. Out of which, three standard charging levels which are predominantly used have been developed. Charging level 1 is based on single phase AC system in which 20 kWh battery is charged in 7 hours, whereas charging level 2 utilizes three phase AC system in which 20 kWh battery is fully charged in 1 hour only. Moreover, charging level 3 also known as fast charging level, the battery gets fully charged in 20-30 minutes [7]. It is based on direct-current (DC) system. It is necessary to have fast charging station for wide adoption of EVs [8]. Charging station (CS) is directly connected to power grid. Equipment's such as transformers, rectifiers are installed to generate dc voltage in the station. Charging connectors are placed in the station and uses DC voltage to charge the batteries in 20-30 minutes. Thus, CS serves electric supply to EV for charging purpose. The easy availability of fast CS is the key to the commercial deployment of EVs [9]. Various issues concerning the development of fast CS have been discussed in the literature, more prominent of them being charging time [7], battery life, accessibility of public CS [10], and integration to the energy supply grid. Additionally, energy consumption and the availability of renewable energy sources for fast CS have also been discussed [11]. Specifically, EVCS based on renewable energy drawn from wind-farm has been described [12]. Further, the recent public sector funding in the commercial deployment of the fast CS in the US and Japan have been studied [13]. It has been shown that a supply of at least 100 kW is necessary for charging a 36 kWh battery-based EVs in 20 minutes. So, for a practical scenario that considers simultaneously charging 10 EVs, the CS would require at least 1000 kW from the electric supply grid, which would incur significant losses [14].

1.1 Related works

Numerous researches have been conducted on the optimal planning of EVCS and its integration to the electric power network. Concerning this, authors have proposed a Spatio-temporal model for analyzing the impact of different EV charging strategies on the electric grid [15]. Thus, for reducing the losses, the CS should be placed nearer to the electric substations. However, the substations are generally located far away from the urban areas, so it increases the energy consumption in the EV in traveling to the CS. Hence, the travel cost and the grid losses both are critical while determining the location of the CS. Moreover, the CS capacity is dictated by the area, increasing the land-area for CS would increase the number of EVs that can be served simultaneously but it also increases the cost associated with development charges and land cost. All these issues show that the placement and sizing of the CS is a non-trivial problem that necessitates a comprehensive study that can address the aforementioned issues [16]. A few of the research work has addressed the placement and sizing problems for the CS [17-24]. In [17], authors have proposed a partitioning-based technique for optimizing the location of the CS by minimizing the traffic loss. An optimal location for the CS has been derived for the city of Lisbon [18]. The location of the CS has been optimized for a minimum station development cost [19]. Optimizing the CS location for the driving pattern of the EVs is focused [20]. Graph theory-based study for optimal placement and sizing of EVCS has been studied [21]. Further, for optimizing the location and size of the CS a two-step-based technique has been proposed [22], whereas particle swarm optimization (PSO) has been used for finding the optimal location [23]. The author utilized Jaya algorithm for the optimal placement of EVCS by optimizing the operating cost, installation cost and power grid loss [24]. A hybrid approach of Chicken swarm optimization (CSO) and Teaching learning based Optimization (TLBO) is utilized for the allocation of EVCS considering various economic and grid operating issues [25]. Several cost functions, including operating cost, investment cost, and maintenance cost, are optimized PSO with taking care of the network operating constraints [26]. The EVCS allocation in the IEEE-123 bus distribution network has been performed using Grey wolf optimization (GWO) without violating system constraints [27]. The optimal placement of EVCS in IEEE-33 bus radial distribution system has been achieved considering uncertainties related to the quantity of EVs to be charged [28]. GWO/WOA has been employed to solve the placement problem. The EVCS planning in distribution network superimposed with road network have been performed. Uncertainties related to EVs are considered using 2m Point Estimation Method. Differential evolution (DE) and Harris Hawks Optimization (HHO) algorithm have been utilized for optimizing the objectives [29]. Optimal placement of EV parking lot have been performed in Beijing district for minimizing the total cost and power loss [30]. Author proposed a two stage approach for the allocation of EV parking lot and renewable energy resources (RESs) in a distribution network [31]. Selection of appropriate bus for EV parking and RESs have been performed considering economic objectives in order to reduce the system loss using genetic algorithm (GA) and PSO algorithm. The results prove the effectiveness of simultaneous allocation of RESs and EV parking lots over the independent allocation of distributed energy resources and EV parking lots. In [32], author proposes the improvements of voltage profile, power loss and loading capabilities of IEEE-69 bus radial distribution system with PV/battery energy storage system (BESS)-powered EV charging stations. An energy management strategy has been introduced to direct the power flow among the EVCS, solar panel, BESS and the utility grid based on time of use of electricity price. Multi-course teaching learning based multi-objective optimization (MCTLBO) has been utilized to optimize the size of PV/BESS system and the locations of CSs in each zone for minimizing the annual CS operating cost and the system active power loss. Allocation of EVCS in presence of capacitors in distribution network is carried out for minimizing the active power losses and maximizing the net profit without violating the operational constraints [33]. Optimal CS planning has been done with traffic constrained framework [34]. Although the CS development is highly prone to public sector policies, not much has been studied regarding the impact of public policy on the development of charging infrastructure. Authors have considered grid loss and cost of construction of CS for optimal sizing and placement of EVCS [35]. GA has been used to solve the optimization problem. Authors utilized PSO in which some modifications in inertia factor are carried out for optimal planning of CS [36]. Construction cost and maintenance cost have been considered as an objective function. The proper placement of EVCS is carried out by considering the cost of installation in traffic network [37]. The cost functions have been modeled for the optimal siting of EVCS assuming traffic and geographical constraints [38]. The main emphasis of this article is to minimize the transportation cost. In

[39], author reduces the power losses by optimal sizing of distributed generation (DG) and optimal siting of EVCS. Initially, DG of optimal size are located to reduce the power losses. In second step, CS are installed optimally for further reduction in power losses. The results are tested in IEEE 33 bus radial distribution network. It is also shown that, this approach of sequential placement of DG and CS introduces addition of new load in the system. Table 1. provides an insight of optimization algorithms used for handling multi-objective cost-based functions.

To the best of author's knowledge, numerous studies have been conducted so far for the siting and sizing of EVCS in distribution and transportation network. But few of them addressed the aforementioned problem by taking the realistic data of EVs into account. Also, impact on reliability of grid in terms of charging cost loss (CCL) is rarely performed in past studies. Reliability evaluation is an important factor to be taken into account for the optimal planning of EVCS. Furthermore, application of intelligent meta-heuristic technique for the particular case study has not been performed in prior works. This research work is focused on optimal planning of EVCS in different well-known areas of South Delhi considering land cost, coordinates i.e., latitude and longitude, elevation and population density of CS location. Moreover, reliability analysis is carried out to investigate the impact on grid. Therefore, this work will be carried out by utilizing grey wolf optimization (GWO) for the planning of CS in the proposed site. GWO is simple to execute because of its simple structure, low storage and computational requirements, good convergence due to continuous search space reduction and less decision variables and its ability to avoid local minima and search for global minima over a large search range.

1.2 Motivations

Availability of charging facilities at public places is the key prerequisite for adoption and rollout of EVs. The charging infrastructure must be smart and can be managed by the government agencies and/or private organizations. The feasibility studies should be carried out before the installation of CS. Services offered aim to enhance the end user customer experience and further promote the ease and convenience of charging EVs. These could include smart end user applications to show location and real-time status of the chargers, simple methods of payment, expert customer support, and service and maintenance of the stations.

The main problem arises when several EVs are plugged into the grid simultaneously which causes grid interruptions. Thus, a proper and robust battery management system is required for managing the CS for charging a considerable number of EVs simultaneously. To adopt the swapping approach, a greater number of batteries are required as a result the cost of the system will increase which is not desirable therefore, charging strategies are to be developed.

As EV adoption increases, therefore it is expected that the CS infrastructure will keep pace with the predicted growth scenarios. However, planning the charging infrastructure is a difficult task because many consumer and local constraints must be taken into account. Some of the factors that will influence EV adoption are the expenses incurred in buying the EV, regulations by the government that encourage the use of EVs, availability of electrical energy, role of renewable energy resources, and concerns about the access and availability of fast-charging stations. Because of all of these factors, it may greatly influence and increase uncertainty to the rate of EV adoption, it remains to be seen whether the expected charging infrastructure will support the demand.

1.3 Contributions

Inspired by the current research in the planning of EVCS, this research work contributes to the overall planning of EVCS around the world, specifically focusing on the different most populated areas of New Delhi. Various intelligent algorithms have been employed to deal with this planning problem. Few of these algorithms search for the optimum solution only in the neighborhood space and do not pay much attention to the global area. Such techniques get trapped to a local optimum solution only. On the other hand, some techniques are good enough to achieve the global solution but cannot explore the local solution. Due to these shortcomings, more efficient algorithms are required for premature concurrence and must have good exploration capability. In this concern, GWO is utilized in this work because of its ability to provide a good balance between the local and global search space.

The main contributions of the proposed work are summarized below:

- This research work emphasizes on the planning of EVCS in most populated areas of South Delhi, India. Allocation of EVCS is based on minimizing the overall cost which includes investment cost, CS electrification cost, EV energy loss cost and travel time cost.
- Reliability evaluation is one of the major concerns for the location planning of EVCS. Hence, charging cost loss (CCL) has been investigated which is considered as an important measure for determining the impact on reliability of grid. CCL refers to the cost of EV which remains uncharged because of improper functioning of grid or grid failure. CCL is evaluated for each CS site for reliability evaluation.
- The proposed algorithm relies on the hunting behavior of grey wolves balancing the exploration and exploitation capability. The performance and stability of the proposed algorithm is confirmed by implementing it on various objective functions considered for the optimal planning of EVCS in most populous areas of South Delhi, India. The supremacy of the proposed technique is verified by comparing the obtained results with other existing methods discussed in literature. The GWO results prove the effectiveness of the proposed technique in determining the global solution in terms of quality of solution and computational burden.

The article is divided into 6 sections. Section 2 focuses on modeling of optimization problem. Constraints and adopted methodology are described in section 3. In section 4, applied intelligent algorithm i.e., GWO and its flow chart is given. Section 5 discusses the results and simulation analysis. Conclusions are provided in section 6.

2. Components and models of electric vehicle charging station

The interaction among CSs, EVs and electric substation is shown in Figure 1 which depicts that CS is connected to electric substation and EV takes power from CS.

Figure 2. shows the single line diagram of interconnection of charging station (m) and substation (x) via feeder (F_m). High voltage and low voltage bus bars of the electric substation are denoted by D6x and D4x respectively. Two transformers are assumed to be connected in parallel in the electric substation. D4i indicates the incoming bus bar of the CS.

2.1 Investment Cost (IC)

The investment cost can be divided into three parts. The first part is the cost incurred for establishing the equipment and facilities for CS. The second part is associated with the rental cost of land and the third part is the cost required for developing the connectors. Investment cost can be formulated mathematically as follows [37]

$$IC = \sum_{i=1}^{N_{CS}} \left(C_{initial} + B.C_{land}.NC_i + C_{con}.CP.\left(NC_i - 1 \right) \right) \tag{1}$$

where, N_{CS} are the number of CS, $C_{initial}$ is the initial fixed cost of CS, B is area requirement per connector in m^2, C_{land} is the rental cost of land for the different sites in areas of South Delhi, NC_i is the number of charging connectors in i^{th} CS, C_{con} is development cost of connector and CP is the rated power of connector.

In the current work, it is assumed that the land is taken on rent for 5 years and hence, the operational and maintenance cost has been ignored. Also, area requirement per connector is considered to be 25 m^2 [41].

The cost of equipment depends on the capacity of CS. The rental cost of land depends on the quality of land and varies in different city locations. In the proposed methodology, CS are installed at locations with different land cost. Since the development cost of the charging connectors will decrease in the future by improving the technology, a considerable portion of the investment cost depends on it. Typical layout of EVCS per connector is shown in Figure 3. For a connector, it is required to have minimum width of 2.74 m and length of 5.28 m as shown in Figure 3. 0.92 minimum clearance of 0.92 m is needed between the two connectors if there is a use of more than one connector.

Moreover, the CS capacity is a function of number of charging connector (NC_i) and rated power of connector (CP).

Thus, the capacity of i^{th}CS can be determined as [42]:

$$CS_c\left(i \right) = CP \times NC_i \tag{2}$$

2.2 CS electrification cost (CSEC)

The overall connection cost strongly depends on the separation between the CS and nearest electric substation as well as the connection technology. The CS and electric substation are assumed to be connected directly by means of dedicated overhead lines. Some conductors which are easily available and frequently used for overhead lines are listed in Table 2. The area of cross-section, rated current for different types of conductors used for overhead lines is also given in Table 2.

The transmission cost of overhead line depends on area of cross-section of line and can be evaluated as follows [35]:

$$CT_i = 8000 + 65.7w \tag{3}$$

where, w denotes the area of cross section of transmission line in mm^2.

The overall connection cost of ith CS can be determined as [37]

$$CSEC = \sum_{i=1}^{N_{CS}} \left(CT_i * D_i \right) \tag{4}$$

D_i represents the distance of ith CS from nearest electric substation in Km. CT_i is the transmission cost of the i^{th}CS and can be expressed in per km. $d_{i,n}$ is a matrix which gives distance of ith CS from n^{th} electric substation in Km. The matrix $d_{i,n}$ is accurately determined using the geographic information.

2.3 EV energy loss cost (EVLC)

For charging the batteries of EV, definite path must be followed by EV to reach the nearest charging station. For j^{th} EV, charging loss can be determined as follows [42]

$$P_{EVL} = \sum_{j=1}^{N_{EV}} \left(SEC * S_j \right) \tag{5}$$

Where SEC is specific electricity consumption of EV in kWh/Km, $S_j$is the distance between the EV and CS in Km and P_{EVL} is the EV energy loss.

Thus, the EV energy loss cost can be calculated as follows:

$$EVLC = P_{EVL} * PE * TD \tag{6}$$

Here, PE denotes the electricity price and TD is the total number of days in 5 years.

In this article, S_j is the distance travelled between EV and CS and it should be evaluated assuming the urban roads. The values of S_j are calculated using geographic information.

2.4 Travel Time Cost (TTC)

Travel time cost is the cost required for reaching the nearest CS from the point where the need of charging arises or point of charging demand. It depends on the distance between the EV position and nearest CS and cost of travelling for per Km of EV. According to [38], it can mathematically be written as :

$$TTC = S_j * C_{EV/Km} \tag{7}$$

Where, TTC is the travel time cost, S_j is the distance between EV position and CS location, $C_{EV/Km}$=cost of traveling of EV per Km.

3. Optimization Problem

This section describes the optimization model based on minimizing the aforementioned objective functions, associated constraints and methodology for assigning the EVs to nearest CS.

3.1 Aim of the optimization

The aim of the optimization problem is the minimization of total cost related to charging demand of EV.

$$MinF = \sum_{i=1}^{N_{CS}} \left(IC_i + CSEC_i \right) + \sum_{j=1}^{N_{EV}} EVLC_j + \sum_{j=1}^{N_{EV}} TTC_j \tag{8}$$

N_{CS} represent the number of CS and N_{EV} indicates the number of EV charged in a day.

3.2 Constraints

The optimization problem for optimal sizing and placement of CS is subjected to some existing constraints associated with CS and EVs which are as follows:

a) Each CS must have at least one charging connector.

$$NC_i \geq 1 \qquad \text{i=1, 2, 3……}N_{CS} \tag{9}$$

(b) The connector must be able to charge all EVs.

$$\sum_{i=1}^{N_{CS}} \left(NC_i \times DE \right) \geq N_{EV} \tag{10}$$

where, DE denotes the maximum number of EVs that can be charged by a connector in one day.

c) The trajectory length of each EV to the CS can be determined as:

$$S_j = \min\left(S_{i,j} \right) \tag{11}$$

where, $s_{i,j}$ is the trajectory length of jth EV to the ith CS.

d) The maximum number of EV that can be charged in one CS is limited by

$$NC_i \times DE \geq N_{EV_i} \tag{12}$$

e) Number of EVs which are charged by each CS is given as follow:

$$N_{EV_i} = \sum_{j=1}^{N_{EV}} \left(1 + \text{sgn}\left(S_j - S_{i,j} \right) \right) / 2 \tag{13}$$

3.3 Assumptions

Various assumptions have been taken in this work which is as follows:

- The positions of EVCS and EV are considered to be distributed normally.
- The number of EVs in a given area decides the number of EVCS in that area.
- EV owner charges their EVs in a fixed CS.

3.4 Methodology

3.4.1 Determination of S and D matrices

Two matrices which represent the distance between substation, CS and EV position and CS need to be modeled. D matrix indicates the distance of each CS from every substation whereas S matrix denotes the distance of each EV from every CS.

Matrix D describes the distance between i^{th} CS and n^{th} electric substation. Hence the order of matrix D would be (i×n).

$$D = \left[d_{i,n} \right]_{N_{CS} \times N_{ES}}$$

$$\text{Where, } d_{i,n} = \sqrt{\left(x_{CS_i} - x_{ES_n} \right)^2 + \left(y_{CS_i} - y_{ES_n} \right)^2} \tag{14}$$

i indicates the number of CS while n denotes the number of electric substations.

$$i = 1, 2, 3 \ldots\ldots\ldots N_{CS}$$

$$n = 1, 2, 3 \ldots\ldots\ldots N_{ES}$$

Matrix S represents the distance between j^{th} EV and i^{th} CS. Thus, the order of matrix L is (j×i).

$$S = \left[S_{i,j} \right]_{N_{EV} \times N_{CS}}$$

$$\text{Where, } S_{i,\mathrm{j}} = \sqrt{\left(x_{CS_i} - x_{EV_j} \right)^2 + \left(y_{CS_i} - y_{EV_j} \right)^2} \tag{15}$$

$$j = 1, 2, 3 \ldots\ldots\ldots\ldots N_{EV}$$

The values of x and y are computed on the basis of geographical information in area under study.

3.4.2 Implementation of the proposed approach

The main goal of optimal siting and sizing of EVCS is to determine where the CS should be installed in order to minimize the investment cost and electrification cost for CS and minimizing energy loss cost and travel time cost for EVs. The flow chart of the proposed approach is shown in Figure 4 and it is based on the following steps.

Step 1: The input data such as fixed cost, connector development cost and other parameters should be defined.

Step 2: The EVCSs are positioned along the main urban roads in South Delhi and EVs locations are generated randomly based on normal distribution functions.

Step 3: Select each charging station and assign EV to each CS based on the criteria of minimum distance.

Step 4: Calculate various objective functions i.e., investment cost, CS electrification cost, EV energy loss cost and travel time cost.

Step 5: Check the subjected constraints using Equation (9)-(13), if satisfied, print the results for each cost functions, else got to step 2.

4. Proposed algorithm for optimization problem

4.1 Brief overview of GWO

Syedali Mirjalili and Lewis [43] in 2014 proposed a novel meta-heuristic based intelligent algorithm named as GWO. Grey Wolf prefers to live in a pack of 5-12 members. They follow a dominant social hierarchy level in which alphas (α) and betas (β) occupy first and second level of hierarchy. Delta (δ) wolves follow alphas and betas. The omega (ω) wolves are the lowest level wolves which act as a scapegoat for the pack. Figure 5. shows the hierarchy of grey wolf.

4.2 Mathematical Modeling of GWO

The major steps involved in the mathematical modeling of GWO include social hierarchy of GWO, encircling prey and hunting prey.

4.2.1 Social hierarchy of GWO

In this algorithm, alpha (α) is considered as the best solution and beta (β) is assumed to be the second-best solution. whereas, delta (δ) is considered as the third best solution and rest of the solutions are considered to be omega (ω).

4.2.2 Grey wolves encircling prey

The hunting process starts with the encircling of prey. The steps of encircling are provided in references [43] as follows:

$$\overrightarrow{M} = \left| \bar{C}.\overrightarrow{Z_p(t)} - \bar{Z}(t) \right| \tag{16}$$

$$\vec{Z}(t+1) = \overrightarrow{Z_p}(t) - N.M \tag{17}$$

Where t indicates the current iteration number, $\vec{N}$ and $\vec{C}$ represents the coefficient vectors, $\vec{Z}$ and $\overrightarrow{Z_P}$ are the position vector of prey and grey wolf respectively and can be determined by the following equations.

$$\overrightarrow{N} = 2n.\overrightarrow{r_1} - a \tag{18}$$

$$\overrightarrow{C} = 2.\bar{r}_2 \tag{19}$$

Where r_1 and r_2 are the random number, whose values lies between 0 and 1 and value of a decreases linearly from 2 to 0 over the course of iterations.

4.2.3 Hunting prey

In order to simulate the hunting process of grey wolves in mathematical form, it is assumed that α, β and δ have better information about the position of the prey. α, β and δ are the best three solutions so far and saved while other search agents update their position at each iteration with respect to their best position. The following mathematical equations [37] were developed in this context.

$$\overrightarrow{M_\alpha} = \left| \vec{C}.\overline{Z}_\alpha - Z(t) \right|;\ \overrightarrow{M_\beta} = \left| \vec{C}.\overline{Z}_\beta - Z(t) \right|;\ \overrightarrow{M_\delta} = \left| \vec{C}.\overline{Z}_\delta - Z(t) \right| \tag{20}$$

$$\overrightarrow{Z_1} = \overrightarrow{Z_\alpha} - \bar{N}.M_\alpha\ ;\ \overrightarrow{Z_2} = \overrightarrow{Z_\beta} - \bar{N}.M_\beta\ ;\ \overrightarrow{Z_3} = \overrightarrow{Z_\delta} - \bar{N}.M_\delta \tag{21}$$

$$\vec{Z}(t+1) = \frac{\overrightarrow{Z_1} + \overrightarrow{Z_2} + Z_3}{3} \tag{22}$$

The GWO's exploration ability represents the search process of grey wolves whereas exploitation ability indicates the attacking process of grey wolves.

4.3 GWO algorithm for optimal siting and sizing of EVCS

Series of steps has been adopted for implementing GWO for minimizing the multi-objective cost-based functions. The procedure is explained with the help of flowchart shown in Figure 6. The initially chosen number of iterations, problem dimension and number of search agents (NSA) are applied.

Step 1: Initialization

The driving source of the algorithm i.e., NSA, max iteration, dimension and boundaries of the problem are initialized. In this study, NSA are assumed to be 30 while maximum number of iterations are taken as 100.

Step 2: Grey wolf positions generation

Population of all search agents is randomly generated using GWO and the first three positions are initialized as alpha, beta and delta. After this, objective function values for each search agent are calculated.

Step 3: Quality solution

The constraints are checked for determining the quality solution. If the constraints are satisfied, then objective function is calculated and if the constraints are violated, the results are discarded.

Step 4: Selecting best positions of search agent

The positions of alpha, beta and delta wolves are updated not including the omega wolf. Equation (20)-(22) are utilized for selecting the best solution so far.

Step 5: Determining the new positions of search agents

The new positions of all search agents are determined and the whole process is repeated.

Step 6: Termination criteria

The termination criteria are set as maximum number of iterations in the proposed work. When the number of iterations exceeds the assigned number of iterations, the simulation will be stopped and the optimized value of objective functions will be displayed.

5. Result and discussion

This section describes the explanation of planning region, reliability analysis, simulation results and main findings.

5.1 Description of system under study

The approach has been applied to an area of 218.5 Km^2 in the South Delhi area of New Delhi, India. The coordinates, elevation and population density of different CS locations are provided in Table 3.

There are 15000 vehicles assumed to be EVs in the study zone. Out of which, only three percent of the total EVs are considered for daily charging. Around 500 EVs have been assumed to be charged every day. EVs positions have been generated randomly based on the Cartesian geographic system as indicated in Figure 7. 11 substations are located nearby 20 CS and provide electric supply to the nearby CS. In Figure 7, green circle, red star and blue diamond represent the EVs, CSs and electric substations considered in this work respectively.

20 CS have been placed in different well-known areas of South Delhi and located on Google Map as shown in Figure 8. In addition to this, input parameters required for solving the objective functions are listed in Table 4.

5.2 Impact on reliability of utility grid

Reliability impact of electric grid is an important factor which must be taken into account for the optimal planning of EVCS [47]. Deterioration of electric substation components such as substation transformer and line causes disruption in providing electric supply to CS. EV charging loss creates issues for CS operators and EV owners in terms of economy. Number of models and reliability indices are needed to address the reliability issue of the utility grid for optimal planning of EVCS.

Availability (ξ) of a basic element can be described in terms of the rate of failure, rate of repair of the basic component of substation i.e., transformer and lines [46]. Thus, the availability (ξ) can be given as:

$$\xi = \frac{\psi}{\psi + \chi} \tag{23}$$

where, χ represents the failure rate (failure per year), Ψ represents the repair rate (repair per year) and r=8760/Ψ is the repair time in hr. Table 5. presents the failure rate and repair rate of substation transformer and lines.

One of the important indexes which measure the reliability impact of utility grid is charging loss cost (CCL). CCL refers to the cost of EVs which remain uncharged due to collapse in the utility grid.

CCL may be defined in terms of operating hours of each CS, CS capacity, unavailability of electric supply for i^{th} CS and uninterrupted charging cost. CCL can be given as follows:

$$CCL(i) = 1825 * \rho(i) * C(i) * \eta_i * ICC \tag{24}$$

Where, 1825 is the number of days in 5 years, ρ(i) is the average operating hours of each CS, η_i denotes the unavailability of electric supply at i^{th} CS and ICC denotes the uninterrupted charging cost in $/kWh.

Average operating hours i.e., ρ(i) of each CS can be determined as follows:

$$\rho(i) = \frac{N_{EV_i}}{NC_i} * T \quad \text{in hours} \tag{25}$$

Where, N_{EV_i} represents the total number of EVs charged at i^{th} CS, NC_i is the number of charging connectors required at i^{th} CS and T is the average charging time of an EV in hour.

ICC refers to the cost paid by CS operator and EV owner due to disruption in EV charging because of collapse in grid.

Unavailability of electric supply for i^{th} CS can be calculated as:

$$\eta_i = 1 - \xi_i \tag{26}$$

$$\xi_i = \xi_{fi} * \xi_{D4x} \tag{27}$$

ξ_{fi} denote the station feeder availability of feeder fi and ξ_{D4x} is the availability of electric supply on bus $D4x$.

Availability of electric supply on bus $D4x$ can be calculated as:

$$\xi_{4x} = \xi_{6x}\left(1 - \left(1 - \xi_T\right)^2\right) \tag{28}$$

where, ξ_T is the availability of power transformer in substation and ξ_{D6x} is the availability of electric supply on bus $D6x$.

Calculation of ξ_{D6x} depends on the nature of substation i.e., whether it is PQ bus or not.

$$\xi_{6x} = \left(1 - \left(1 - \xi_T\right)^2\right) \tag{29}$$

5.3 Simulation results and main achievements

Optimized value of objective function has strong dependence on CS capacity, number of EVs charged at each CS and number of charging connectors utilized to recharge the EVs. For charging, EV will follow the path of minimum distance to reach the CS. Number of EVs charged at each CS can be approximately determined by Equation (13). Figure 9. shows the number of EV assigned at each CS.

Thus, the number of charging connectors installed at each CS can be obtained as follows:

$$NC_i = \frac{N_{EV_i} \times NC(\max)}{N_{EV}} \tag{30}$$

where, N_{EV_i} represents number of EVs charged at CS(i)and NC(max) denotes the maximum limit of charging connectors. Thus, number of charging connectors utilized at each CS for feeding assigned number EVs are displayed in Table 6.

Based on this, different cost functions considered in this paper can be evaluated.

Investment cost mainly depends on land cost and number of charging connectors. Areas in Delhi are divided from category A to category H on the basis of land cost. In the considered study zone, different areas come under different category. Land cost of different considered areas are indicated in Table 7.
Investment cost for different CS sites have been evaluated using the proposed GWO. The overall investment cost incurred by utilizing GWO comes out to be 8657600 $. The effectiveness of the proposed GWO is confirmed by comparing the results with other intelligent techniques such as PSO (9143400 $). It is revealed that the investment cost calculated using GWO is 5.3% less than that of PSO. This comparison shows the preeminence of GWO over PSO. In this regard, the comparative analysis of investment cost investigated using different intelligent algorithms is shown in Figure 10. It should also be remembered that investment costs are high in areas with high land costs and that use a large number of charging connectors. CS that need a greater number of connectors have a higher set-up cost, which leads to a substantial increase in investment costs because connector set-up costs account for a significant portion of the total cost. That's why areas such as New Friends Colony, Vasant Vihar, Chanakyapuri and Golf Links have high investment cost because of their high land cost and ample quantity of connectors required to serve EVs in these regions. On the other hand, Saket, Hauz Khas and Kalkaji etc., have lower investment costs due to lower land costs and a limited number of charging connectors.

Comparison of investment cost of each CS site obtained using GWO and PSO is presented in Table 8. Further improvement in investment cost can be made when the government and public sector unit contributes to the development of charging station and utilizes available public lands for its installation.

In addition, CS electrification cost is highly influenced by the separation between the CS and electric substation. However, distance of each CS from every substation have been evaluated on the basis of Cartesian geographic information. Based on this, electrification cost of each CS site is determined by using proposed GWO. CS electrification cost obtained using GWO and PSO are displayed in Table 8. The performance of GWO technique is proven by comparing the attained results with those of PSO. Application of GWO results in noticeable improvement in CS electrification cost by 5.2% less than that of PSO. Figure 11 shows the comparison of CS electrification cost obtained using GWO and PSO. The greater the distance between the CS and the electric substation, the higher the cost of CS electrification, and vice versa. Table 8 shows that CS in Safdarjung, Chanakyapuri, and Vasant Kunj have high CS electrification costs due to their remote location from the electric substation whereas because of the close proximity to the substation, the Chirag Delhi and Green Park charging station sites have low costs. As a result, in order to meet lower CS electrification costs, the CS operator must choose the location wisely. Even, when the electric utility is obligated to electrify the stations at no cost to the station owners, the cost of CS electrification can be reduced.
Furthermore, EV energy loss cost or cost of energy consumption of EVs relies on the distance between the EV position and CS. CS in the neighborhood of electric substation may be at a far distance from the EV position. Hence, it results in increase in energy loss cost. Distance between each EV (i.e., 500 EVs in this work) and CS is calculated based on geographic information. EVs are assigned to the nearest charging station based on minimum distance criterion. In this way, energy loss cost of 500 EVs have been calculated by optimizing the distance to nearby charge stations. The proposed GWO leads to considerable reduction in energy loss cost i.e., 15.5% less as compared to that of PSO. The obtained results using GWO prove its superiority over other algorithms. The varying behavior of EV energy loss cost obtained using algorithms considered in this work are shown in Figure 12. It has been reported that the smaller the distance between the EV location and the CS site, the lower the EV's energy consumption, and vice versa. If the needs of EV owners are neglected, the cost of EV energy loss may be reduced.
Travel time cost depends on the distance between the charging station and the location where need of charging EV arises, as well as the cost of travelling of EV per Km. Lesser the distance between the EV and CS, smaller will be the travel time cost paid by EV owner. Hence, travel time cost of all 500 EVs to reach the nearby charging station have been calculated using proposed GWO and compared with PSO to prove its dominance. It is observed that, GWO results in significant reduction in travel time cost as compared to the results attained via PSO. GWO provides optimized value of travel time cost i.e., 280.43 $ which is 9.1% less than that of PSO (308.72 $). This proves the supremacy of proposed GWO algorithms over PSO. The travel time cost of each EV (i.e., 500 EVs) determined using GWO and its comparison with PSO is portrayed in Figure 13. The cost of travel time can be further decreased by placing more charging stations along EV routes, as this reduces the distance that EVs would travel for charging.

In addition to this, impact on reliability of grid is investigated in terms of CCL. The availability of electric components in electric substation is calculated using the Equation (26) -(29). CS capacity, effective operating hours and electric supply availability at each CS site are presented in Table 9.

CCL differ for each CS as it depends on CS capacity, effective operating hours and electric supply availability. The optimized values of CCL for each CS site obtained by applying GWO and its comparison with PSO is shown in Figure 14.

Equation (26) -(29) introduces the impact on reliability of grid in optimal planning of EVCS. Basic parameters required for the evaluation of reliability are listed in Table 5. Comparison of the value of CCL for each CS site obtained using GWO and PSO considering the impact on reliability of grid network are shown in Table 10. Also, values of CCL are much dependent on availability of electric supply at particular CS as shown in Table 10. CCL value is high for those CS where percentage of unavailability of electric supply is high and vice versa.

It can be realized in Table 5. that ICC is considered as 10 times the electricity price to reveal the loss due to the power outage.

The objective functions represented by Equation (1) -(7) and set of technical constraints i.e., Equation (9) -(13) have been solved using grey wolf optimization.

The optimized value of all the objective functions considered for the optimal planning of EVCS and their impact on reliability of grid is summarized in Table 11. It can be seen in Table 11; CCL is smaller as compared to the other cost functions. Though, CCL can be of significant value when EV loss is ignored in the optimization problem.

6. Conclusion

In this paper, a new approach for solving the multi-objective optimization problems for optimal siting and sizing of EVCS in different areas in South Delhi, New Delhi, India is proposed. Grey wolf optimization has been utilized to solve the multi-objective cost functions and the obtained results are compared with particle swarm optimization for validation purpose. The cost function includes investment cost, CS electrification cost, EV energy loss cost and travel time cost. Investment cost and CS electrification cost of all CSs have been calculated and found that the investment cost strongly depends on the land cost and number of connectors used in each CS. While, CS electrification cost depends on the distance between the CS and electric substation. On the other hand, EV energy loss cost and travel time cost have been evaluated for all EVs. It is revealed that travel time cost and EV energy loss cost depend on the distance between the CS and EV location. After comparing the results obtained using GWO and PSO, it is revealed that GWO is more effective than PSO for the considered case study. Furthermore, reliability indices i.e., CCL is an important index which gives an idea about the reliability of a system. CCL is a function of CS capacity, effective working hours of CS and unavailability of electric supply at each CS. CCL has been determined for each CS and results reveal that CCL will be high for CS having less availability of electric supply and vice versa.

References

1. Khooban, M.H. Sha-Sadeghi, M. and Niknam, T. “Analysis, control and design of speed control of electric vehicles delayed model: multi-objective fuzzy fractional-order PIλDμ controller”, *IET Sci. Meas. Technol*., **11**(3), pp. 249–261 (2007).

2. Khooban, M.H. Niknam, T. Blaabjerg, F. “Free chattering hybrid sliding mode control for a class of non-linear systems: electric vehicles as a case study”, *IET Sci. Meas. Technol*., **10**(7), pp. 776–785 (2016).

3. Saxena, S. Gopal, A. and Phadke, A. “Electrical consumption of two-, three- and four-wheel light-duty electric vehicles in India”, *Appl Energy*., **115**, pp. 582-590 (2013).

4. Saarenpää, J. Kolehmainen, M. and Niska, H. “Geodemographic analysis and estimation of early plug-in hybrid electric vehicle adoption”, *Appl Energy*., **107**, pp. 456–464 (2013).

5. Moradi, M.H. Abedini, M. Tousi, S.R. “Optimal siting and sizing of renewable energy sources and charging stations simultaneously based on differential evolution algorithm”, *Int. J. Electr. Power Energy Syst.,* **73**, pp. 1015-1024 (2015).

6. Davidov, S. and Pantoš, M. “Planning of electric vehicle infrastructure based on charging reliability and quality of service”, *Energy*, **118**, pp. 1156–1167 (2017).

7. Sheikhi, A. Bahrami, S. Ranjbar, A.M., et al. “Strategic charging method for plugged in hybrid electric vehicles in smart grids; a game theoretic approach”, *Int J Electr Power Energy Syst*., **53**, pp. 499–506 (2013).

8. Nansai, K. Tohno, S. Kono, M., et al. “Life-cycle analysis of charging infrastructure for electric vehicles”, *Appl Energy*., **70**, pp. 251–265 (2001).

9. Fox, G.H. “Electric vehicle charging stations: are we prepared?”, *IEEE Ind Appl Mag*., **19**, pp. 32–38 (2013).

10. Schroeder, A. and Traber, T. “The economics of fast charging infrastructure for electric vehicles”, *Energy Policy*., **43**, pp. 136–144 (2012).

11. Anegawa, T. “Desirable characteristics of public quick charger”, (2009).

12. Weis, A. Jaramillo, P. and Michalek, J. “Estimating the potential of controlled plug-in hybrid electric vehicle charging to reduce operational and capacity expansion costs for electric power systems with high wind penetration”, *Appl Energy*., **115**, pp. 190–204 (2014).

13. Brown, S. Pyke, D. and Steenhof, P. “Electric vehicles: the role and importance of standards in an emerging market”, *Energy Policy*., **38**, pp. 3797–3806 (2010).

14. Schey, S. Scoffield, D. and Smart, J. “A first look at the impact of electric vehicle charging on the electric grid in the EV project”, *World electric vehicle journal*., **5**(3), pp. 1–12 (2012).

15. Mu, Y. Wu, J. Jenkins, N., et al. “A spatial-temporal model for grid impact analysis of plug-in electric vehicles”, *Appl Energy*., **114**, pp. 456–465 (2014).

16. Diaz-Chavez, R. Woods, J. San Román, T.G., et al. “Regulatory framework and business models for charging plug-in electric vehicles: infrastructure, agents, and commercial relationships”, *Energy Policy*., **39**, pp. 6360–6375 (2011).

17. Ge, S. Feng, L. and Liu, H. “The planning of electric vehicle charging station based on Grid partition method”, *In: 2011 Int. conf. electr. control eng., IEEE*, pp. 2726–2730 (2011).

18. Frade, I. Ribeiro, A. Antunes, A.P., et al. “An optimization model for locating electric vehicle charging stations in central urban area”, *In: Transp. res.board 90th annu. meet.,* Washington, D.C., **2252**, pp. 91-98 (2011).

19. Li, Y. Li, L. Yong, J., et al. “Layout planning of electrical vehicle charging stations based on genetic algorithm”, *Lect Notes Electr Eng.*, **99**, pp. 661–668 (2011).

20. Hisatomo Hanabusa, R.H. “A study of the analytical method for the location planning of charging stations for electric vehicles”, pp. 596–605 (2011).

21. Jia, L. Hu, Z. Song, Y., et al. “Optimal siting and sizing of electric vehicle charging stations”, *In: 2012 IEEE int. electr. veh. conf., IEEE*, pp. 1–6 (2012).

22. Liu, Z. Wen, F. and Ledwich, G. “Optimal planning of electric-vehicle charging stations in distribution systems”, *IEEE Trans Power Deliv*., **28**, pp. 102–110 (2013).

23. Liu, Z. Zhang, W. Ji, X., et al. “Optimal planning of charging station for electric vehicle based on particle swarm optimization”, *IEEE PES Innov Smart Grid Technol IEEE*., pp. 1–5 (2012).

24. Ajit, KM and Suresh, B. “Optimal placement of electric vehicle charging station using jaya algorithm”, *Recent advances in power system*, pp. 259-266 (2020).

25. Sanchari, Deb. Kari, T. Karuna, K., et al. “Charging station placement for electric vehicles: a case study of Guwahati City, India”, *IEEE Access*, **7**, pp. 100270-100282 (2019).

26. Han, C. Jin, N. and Bichao, Y. “PSO-based siting and sizing of electric vehicle charging stations,” *Journal of Physics: Conference Series*, **1346**, pp. 1-7 (2019).

27. Akanksha. S, Kusum, V. and Rajesh, K. “Multi-objective synergistic planning of EV fast charging stations in the

distribution system coupled with the transportation network", *IET Gener. Transm. Distrib.,* **13**(15), pp. 3421-3432 (2019).

28. Arnab, P. Aniruddha, B. and Chakraborty, AK. "Allocation of EV Fast Charging Station with V2G Facility in Distribution Network", *8th International Conference on Power Systems (ICPS),* pp.1-6 (2019).

29. Arnab, P. Aniruddha, B. and Chakraborty, AK. "Allocation of electric vehicle charging station considering uncertainties", *Sustainable Energy, Grids and Networks*, pp. 1-28 (2020).

30. Zhang, Y. Zhang, Q. Farnoosh, A., et al. "GIS-Based Multi-Objective Particle Swarm Optimization of Charging Stations for Electric Vehicles", *Energy*, **169**, pp. 844-853 (2019).

31. Amini, M. H., Moghaddam, M. P., and Karabasoglu, O. "Simultaneous allocation of electric vehicles' parking lots and distributed renewable resources in smart power distribution networks", *Sustain. Cities Soc.*, **28**, pp. 332-342 (2017).

32. Kumari, K. Manas, N. and Chinmay, N. "PV/BESS to support electric vehicle charging station integration in a capacity constrained power distribution grid using MCTLBO", *Scientia Iranica*, pp. 1-41 (2020). DOI:10.24200/SCI.2020.5128.1112

33. Mohd B and Mohammad R, "Integration of electric vehicle charging stations and capacitors in distribution systems with vehicle-to-grid facility," *Energy Sources, Part A Recover. Util. Environ. Eff.*, pp. 1–30 (2021). DOI: https://doi.org/10.1080/15567036.2021.1923870

34. Wang, G. Xu, Z. Wen, F., et al. "Traffic-constrained multi objective planning of electric-vehicle charging stations", *IEEE Trans Power Deliv*., **28**, pp. 2363–2372 (2013).

35. Sadeghi-Barzani, P. Rajabi-Ghahnavieh, A. and Kazemi-Karegar, H. "Optimal fast charging station placing and sizing", *Appl. Energy*., **125**, pp. 289– 299 (2014).

36. Phonrattanasak, L. Nopbhorn, P. "Optimal location of fast charging station on residential distribution grid," *Int. J. Innov. Manag. Technol*, **3**(6), pp. 675–681 (2012).

37. Hamid, S. Hasan, D. Hadi, R. "Cost-based optimal siting and sizing of electric vehicle charging stations considering demand response programmes", *IET Gener. Transm. Distrib*., **12**(8), pp. 1712–1721 (2018).

38. Sanchari, D. Karuna, K. Xiao-Zhi, G., et al. "Optimal placement of charging station using CSO-TLBO algorithm", *2017 International conference on Research in Computational Intelligence and Communication Networks (ICRCICN),* pp. 84-89 (2017).

39. Jamian, J.J. Mustafa, M.W. Mokhlis, H. "Minimization of power losses in Distribution System via Sequential Placement of Distributed Generation and Charging Station", *Arab J Sci Eng*., **39**, pp. 3023-3031 (2014).

40. Mohd, B. and Mohammad R. "Electric vehicles in a smart grid: a comprehensive survey on optimal location of charging station", *IET Smart grid*, **3(**3), pp. 267-279 (2020).

41. Awasthi, A. Venkitusamy, K. Padmanaban, S., et al. "Optimal planning of electric vehicle charging station at the distribution system using hybrid optimization algorithm," *Energy*, **133**, pp. 70–78 (2017).

42. Islam, M. M. Shareef, H. and Mohamed, A. "Optimal siting and sizing of rapid charging station for electric vehicles considering Bangi city road network in Malaysia," *Turkish J. Electr. Eng. Comput. Sci.*, **24**, pp. 3933–3948 (2016).

43. Seyedali, M. and Andrew, L. "Grey Wolf Optimizer", *Adv. Eng. Softw*., **69**, pp. 46-61 (2014).

44. "https://www.google.com/maps/d/u/0/edit?mid=1MYBWdZhX_Ay0IWXwRHNjbXcigF88."

45. "https://www.google.com/search?q=latitudes+and+longitudes+of+areas+of+south+delhi&oq=latitudes+and+longitudes+of+areas+of+south+delhi&aqs=chrome..69i57j33i22i29i30.13439j0j7&sourceid=chrome&ie=UTF-8."

46. Allan, R.N. Billinton. “Reliability evaluation of power systems”, Springer (1996).
47. Hamedani-Golshan, M. “Planning & calculations of electrical distribution systems,” *3rd ed. Isfahan Islam. Azad Univ. Majlesi Branch*, (2007).
48. “https://www.myloancare.in/delhi-circle-rate-revised.”

Abbreviations and symbols:

IC	Investment cost	**D**	Distance between i^{th}CS and nearest electric substation
CSEC	Charging station electrification cost	$\mathbf{P_{EVL}}$	EV energy loss
EVLC	Electric vehicle energy loss cost	**DE**	Maximum number of EVs that can be charged by a connector
TTC	Travel time cost	$\mathbf{C_{EV/Km}}$	Cost of traveling of EV per Km
CCL	Charging cost loss	**ES**	Electric substation
SEC	Specific electricity consumption of EV	$\mathbf{N_{ES}}$	Number of electric substations
PE	Electricity price	$\mathbf{x_{CS}}$	Abscissae of charging station
TD	Total number of days in 5 years	$\mathbf{y_{CS}}$	Ordinate of charging station
ICC	Uninterrupted charging cost	$\mathbf{x_{EV}}$	Abscissae of electric vehicle
$\mathbf{C_{initial}}$	Initial fixed cost of charging station	$\mathbf{y_{EV}}$	Ordinate of electric vehicle
$\mathbf{C_{land}}$	Rental cost of land	$\mathbf{x_{ES}}$	Abscissae of electric substation
B	Area requirement per connector	$\mathbf{y_{ES}}$	Ordinate of electric substation
$\mathbf{N_{CS}}$	Number of charging station	$\boldsymbol{\xi}$	Availability of electric supply
$\mathbf{N_{EV}}$	Number of electric vehicles	$\boldsymbol{\chi}$	Failure rate
i	Charging station index	$\boldsymbol{\Psi}$	Repair rate
j	Electric vehicle index	**r**	Repair time
$\mathbf{NC_i}$	Number of charging connectors at i^{th}CS	$\boldsymbol{\rho}(\mathbf{i})$	Average operating hours of i^{th}CS
CP	Rated power of connectors	$\boldsymbol{\eta_i}$	Unavailability of electric supply at i^{th}CS
$\mathbf{CS_C}(\mathbf{i})$	Capacity of i^{th}CS	**T**	Average charging time of an EV
w	Area of cross section of transmission line	$\boldsymbol{\xi_{fi}}$	Station feeder availability of feeder fi
$\mathbf{CT_i}$	Transmission cost of i^{th}CS	$\boldsymbol{\xi_{D4x}}$	Availability of electric supply on bus $D4x$.
$\mathbf{S_j}$	Trajectory length to the charging station	$\boldsymbol{\xi_T}$	Availability of power transformer in substation
$\mathbf{s_{i,j}}$	Trajectory length of jth EV to i^{th}CS	$\boldsymbol{\xi_{D6x}}$	Availability of electric supply on bus $D6x$

List of Figure Captions

Figure 1. Interaction among CSs, EVs and electric substation
Figure 2. Single line diagram for station connection to electric grid [35]
Figure 3. General CS station layout per connector [42]
Figure 4. Flow chart of the proposed approach
Figure 5. Social hierarchy in grey wolves
Figure 6. Flowchart of GWO algorithm for the proposed approach
Figure 7. Positions of EVs, charging stations and electric substations
Figure 8. Location of CS in South Delhi Areas in Google Map [44]
Figure 9. Number of EVs charged at each CS
Figure 10. Investment Cost for different CS in South Delhi areas
Figure 11. CS electrification cost for different CS in South Delhi area
Figure 12. EV energy loss cost for 500 EVs
Figure 13. Travel time cost for 500 EVs
Figure 14. Charging cost loss for CS installed in South Delhi areas

List of Table Captions

Table 1. Analysis of various intelligent algorithms applied for handling cost-based objective functions [40]
Table 2. Conductor used for dedicated overhead line [45]
Table 3. Charging Station Information [45]
Table 4. Input Parameter Required for objective function calculation [37, 38]
Table 5. Data required for determining the reliability [46]
Table 6. Optimized number of charging connectors required at different CS installed in South Delhi areas
Table 7. Land cost of areas in South Delhi [48]
Table 8. Investment cost and CS electrification cost of each CS installed in South Delhi areas
Table 9. Average working hours, unavailability of electricity and capacity of each CS
Table 10. Charging cost loss (CCL) of each charging station
Table 11. Optimized values of objective functions

Figures

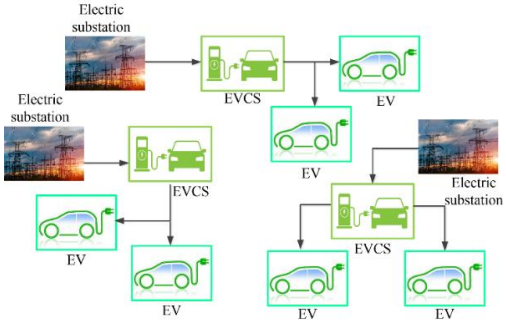


Figure 1

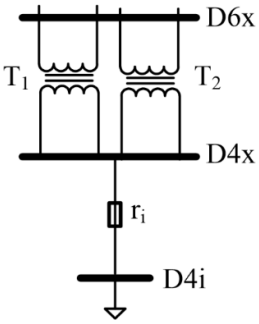


Figure 2

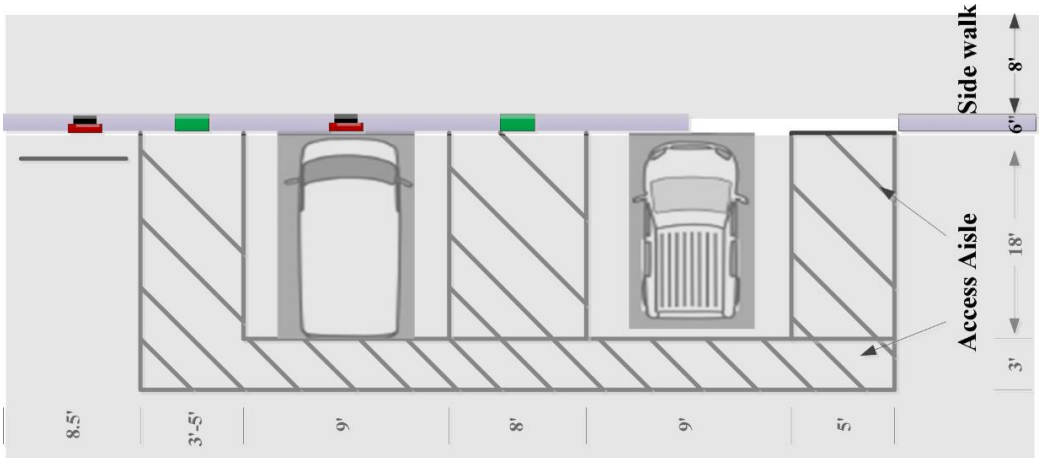


Figure 3

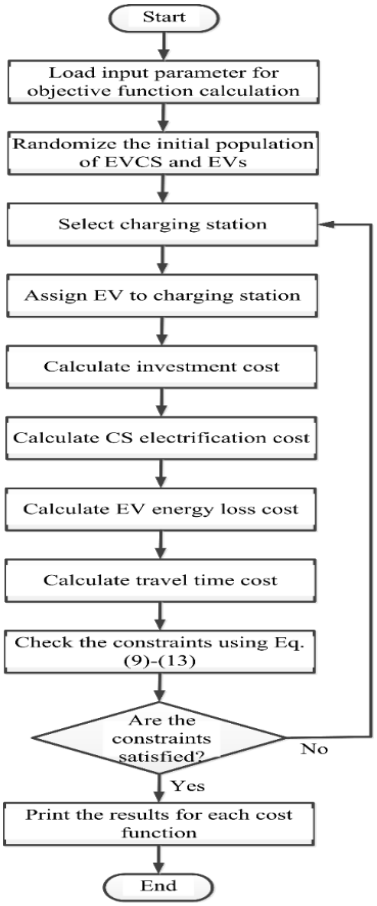


Figure 4

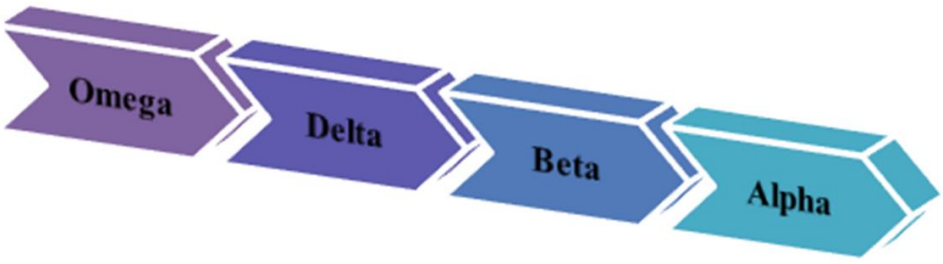


Figure 5

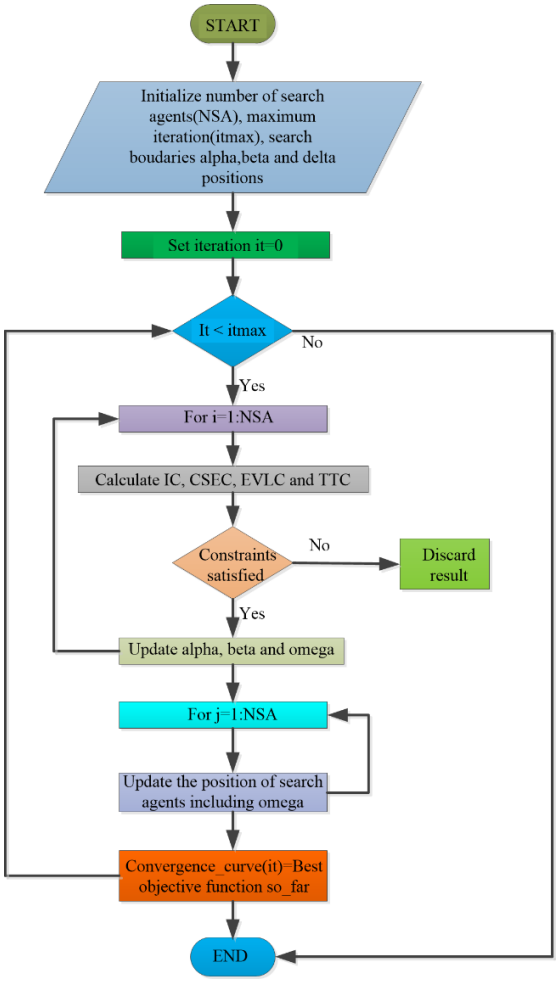


Figure 6

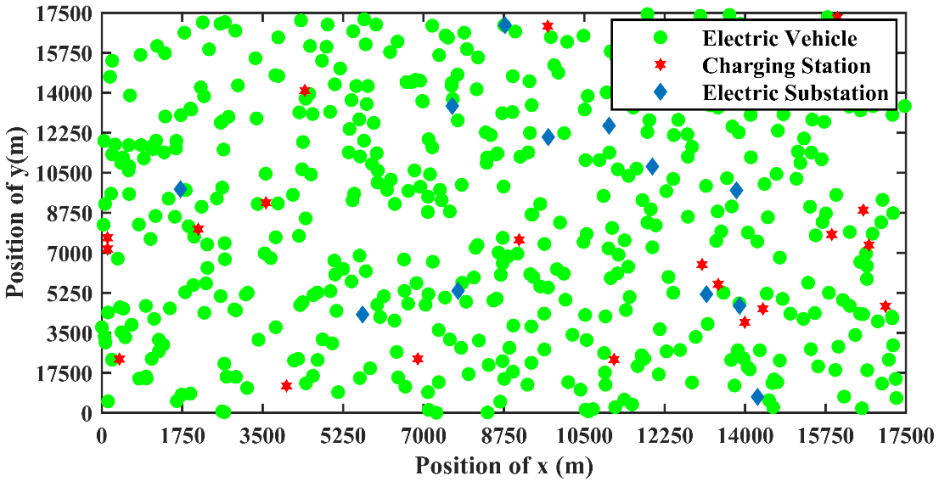


Figure 7

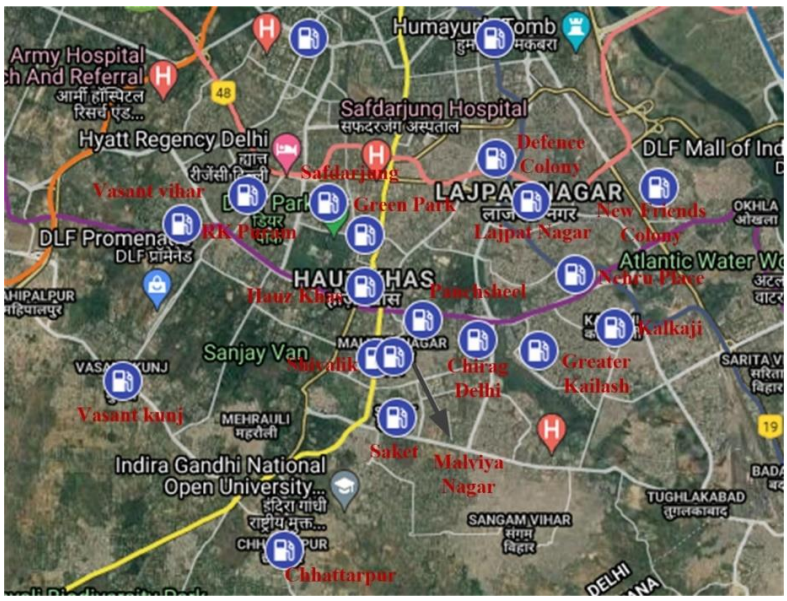


Figure 8

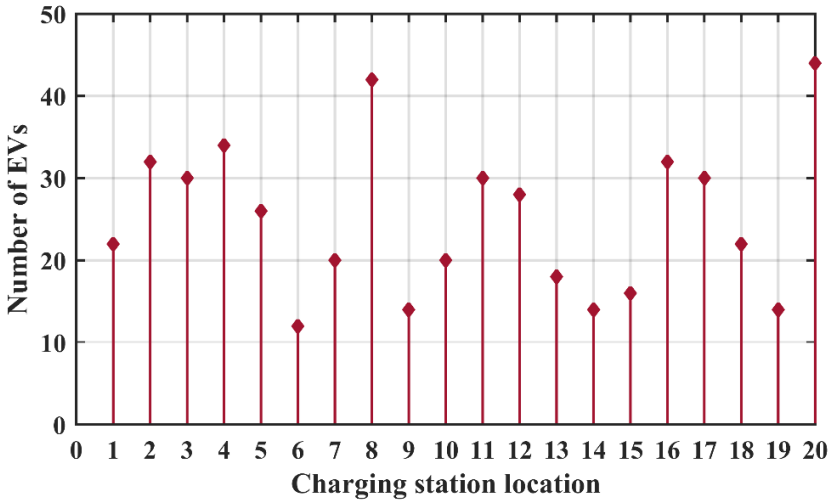


Figure 9

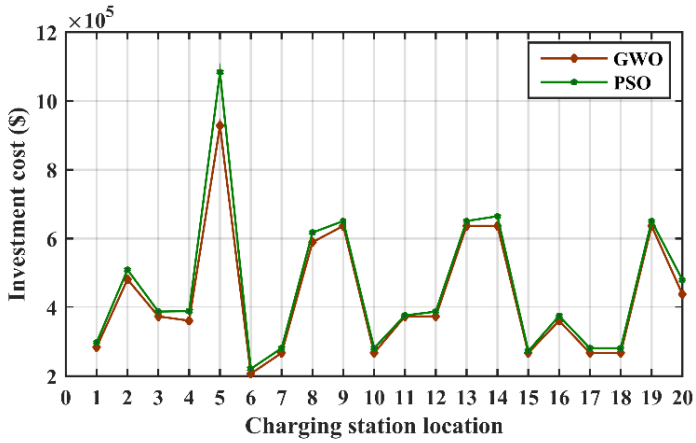


Figure 10

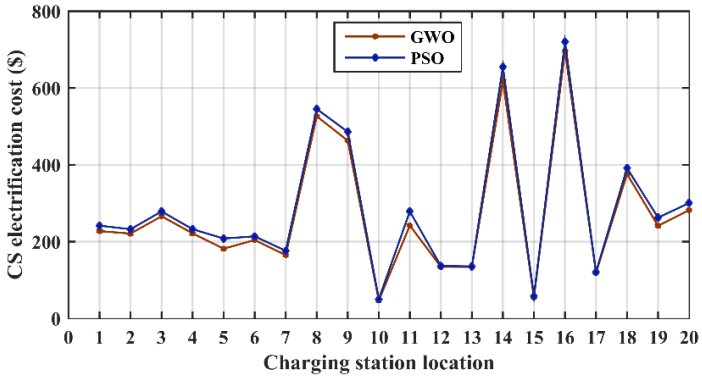


Figure 11

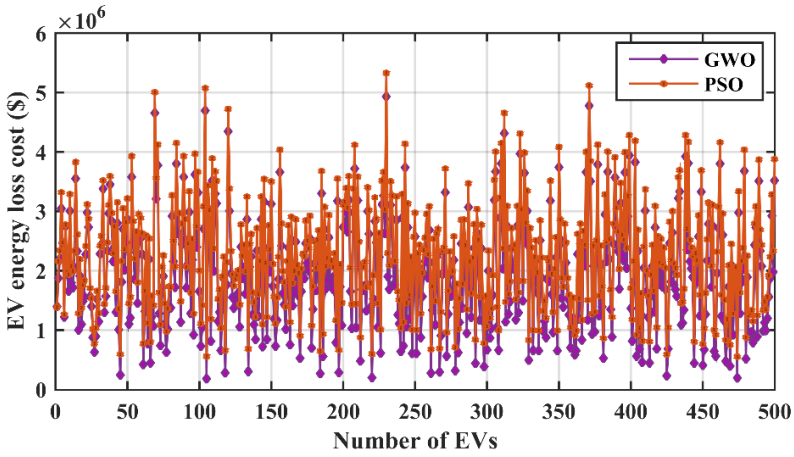


Figure 12

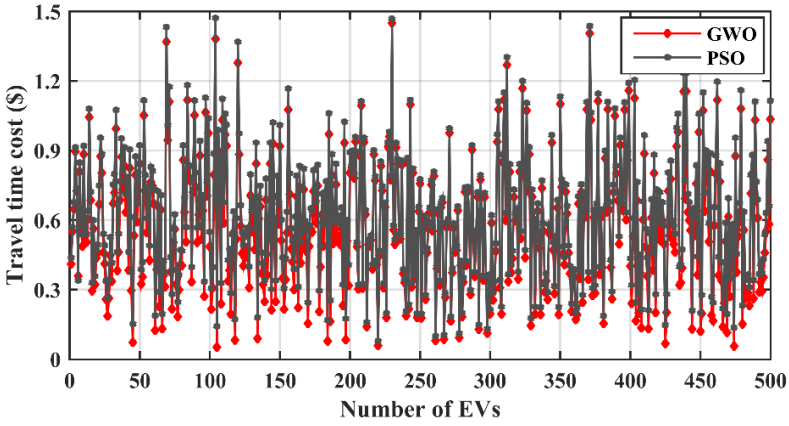


Figure 13

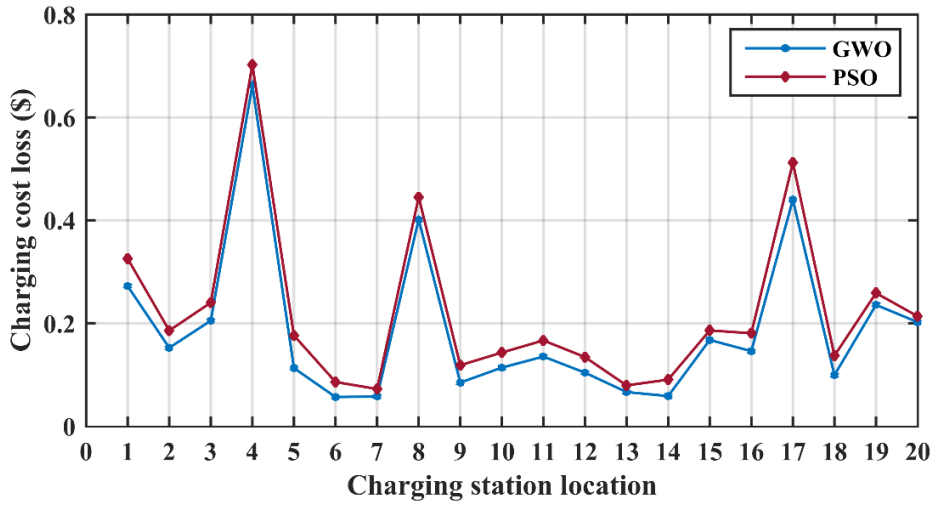


Figure 14

Tables

Table 1

[Ref]	Cost functions	Algorithm	Category	Attributes
[36]	Minimizing investment cost	ACO	Swarm Intelligent	Exploitation capability is weak
[37]	Minimizing connection cost	PSO	Swarm Intelligent	Processing time is more for multi-modal problems.
[37]	Minimizing EV energy loss cost	PSO	Swarm Intelligent	Depending on own parameters such as inertia weight, social parameters and cognitive term.
[35]	Minimizing EV energy loss cost	GA	Evolutionary	Issue in tuning the parameter due to complex mutation, speed of convergence and crossover operator.
[38]	Minimizing installation cost	TLBO	Population-based optimization	Memory required for computation purpose is less
[38]	Minimizing travel time cost	TLBO	Population-based optimization	Initially, it could not solve problems continuous in nature

Table 2

Name	Cross section area in mm^2	Rated current in ampere
FOX	42.77	192
MINK	73.6	288
DOG	118.5	380
PARTRIDGE	156.9	460

Table 3

Charging Station Location	Coordinates		Elevation(m)	Population density (people per Km^2)
	Latitude	Longitude		
Saket	28.5221	77.2012	240	20110
Shivalik	28.5340	77.2053	227	19557
Greater Kailash	28.5555	77.2337	224	26450
Lajpat Nagar	28.5649	77.2403	211	34599
New Friends Colony	28.5675	77.2691	208	8956
Kalkaji	28.5400	77.2592	239	24803
Hauz Khas	28.5479	77.2031	223	7974
Safdarjung	28.5647	77.1949	221	29682
Vasant Vihar	28.5603	77.1617	240	17475
Green Park	28.5584	77.2029	218	35903
Panchsheel	28.5415	77.2161	224	11576
Defence Colony	28.5734	77.2326	212	16837
Nehru Place	28.5503	77.2502	231	24036
Chanakyapuri	28.5972	77.1904	220	7498
Chirag Delhi	28.5376	77.2283	225	22552
Vasant Kunj	28.5293	77.1484	264	10536
Chhatarpur	28.4959	77.1848	261	13101
RK puram	28.5660	77.1767	229	14620
Golf Links	28.5973	77.2323	209	5919
Malviya Nagar	28.5342	77.2094	226	29945

Table 4

Parameter	Value	Unit
N_{EV}	500	−
N_{CS}	20	−
N_S	11	−
SEC	7	Km/kWh
PE	90.48	$
$\mathbf{C_{initial}}$	72204.52	$
$\mathbf{C_{con}}$	214.89	$/kW
CP	96	kW
DE	30	−
TD	1825	Days
w	156.9	mm^2
T	0.5	hr
$C_{EV/Km}$	0.34	$/Km

Table 5

		Value	Unit
Line	Failure rate	0.1	Failure/Km/year
	Repair time	4	hour
Transformer	Failure rate	0.1	Failure/year
	Repair time	100	hour
ICC		0.90	$/kWh

Table 6

Charging Station Location	**Number of charging connectors utilized**
Saket	3
Shivalik	4
Greater Kailash	3
Lajpat Nagar	4
New Friends Colony	3
Kalkaji	2
Hauz Khas	2
Safdarjung	5
Vasant Vihar	2
Green Park	2
Panchsheel	3
Defence Colony	3
Nehru Place	2
Chanakyapuri	2
Chirag Delhi	2
Vasant Kunj	4
Chhatarpur	3
RK puram	3
Golf Links	2
Malviya Nagar	5

Table 7

Charging Station Location	Land Cost ($/m^2)
Saket	2178.77
Shivalik	3349.87
Greater Kailash	3349.87
Lajpat Nagar	2178.77
New Friends Colony	10539.82
Kalkaji	2178.77
Hauz Khas	3349.87
Safdarjung	3349.87
Vasant Vihar	10539.82
Green Park	3349.87
Panchsheel	3349.87
Defence Colony	3349.87
Nehru Place	10539.82
Chanakyapuri	10539.82
Chirag Delhi	3349.87
Vasant Kunj	2178.77
Chhatarpur	3349.87
RK puram	3349.87
Golf Links	10539.82
Malviya Nagar	2178.77

Table 8

Charging Station Location	Investment cost ($)		CS electrification cost ($)	
	GWO	PSO	GWO	PSO
Saket	284200	298200	227.36	241.5
Shivalik	481600	509600	220.864	232.218
Greater Kailash	373800	387800	266	278.572
Lajpat Nagar	361200	389200	221.424	232.246
New Friends Colony	928200	1083600	181.44	208.026
Kalkaji	207200	221200	204.064	213.696
Hauz Khas	267400	281400	164.752	176.092
Safdarjung	589400	617400	526.288	545.398
Vasant Vihar	637000	651000	462.896	485.8
Green Park	267400	281400	46.704	50.05
Panchsheel	373800	376600	242.032	278.558
Defence Colony	373800	387800	134.4	136.556
Nehru Place	637000	651000	134.064	135.492
Chanakyapuri	637000	665000	618.24	654.92
Chirag Delhi	267400	273000	54.656	57.988
Vasant Kunj	361200	375200	696.416	720.398
Chhatarpur	267400	281400	118.832	120.568
RK puram	267400	281400	375.872	391.384
Golf Links	637000	651000	241.136	262.43
Malviya Nagar	438200	480200	281.792	300.412
Total	**8657600**	**9143400**	**5419.232**	**5722.304**

Table 9

Charging Station Capacity (kW)	Effective Operating Hours of each CS, $\rho(i)$ in hr	Unavailability of electric supply (ξ_i) ($\times 10^{-3}$)
288	3.66	0.2750
384	4	0.1060
288	5	0.1523
384	4.25	0.4343
288	4.33	0.0967
192	3	0.1050
192	5	0.0640
480	4.2	0.2131
192	3.5	0.1347
192	5	0.1267
288	5	0.1004
288	4.66	0.0828
192	4.5	0.0820
192	3.5	0.0929
192	4	0.2333
384	4	0.1015
288	5	0.3271
288	3.66	0.1006
192	3.5	0.3752
480	4.4	0.1024

Table 10

Charging Station Location	Charging Cost Loss ($)	
	GWO	**PSO**
Saket	0.272104	0.3255
Shivalik	0.152124	0.18564
Greater Kailash	0.204904	0.23996
Lajpat Nagar	0.662242	0.7021
New Friends Colony	0.112672	0.17612
Kalkaji	0.056504	0.08568
Hauz Khas	0.057484	0.0721
Safdarjung	0.401408	0.44492
Vasant Vihar	0.084574	0.1183
Green Park	0.113638	0.14294
Panchsheel	0.135086	0.16646
Defence Colony	0.103824	0.13412
Nehru Place	0.066192	0.0791
Chanakyapuri	0.058324	0.0903
Chirag Delhi	0.167412	0.18592
Vasant Kunj	0.14567	0.18046
Chhatarpur	0.440104	0.51184
RK puram	0.099078	0.13692
Golf Links	0.235578	0.25858
Malviya Nagar	0.202062	0.2135
Total	**3.770984**	**4.45046**

Table 11

Objective Functions	Cost ($)	
	GWO	**PSO**
Investment cost	8657600	9143400
CS electrification cost	5419.232	5722.304
EV energy loss cost	9.51×10^{8}	11.26×10^{8}
Travel Time Cost	280.4305	308.7251
Charging cost loss	3.770984	4.45046

Author's Biography:

Mohd Bilal received his B.Tech. degree in electrical engineering and M. Tech. degree in instrumentation and control engineering from Aligarh Muslim University, Aligarh, India, in 2014 and 2016, respectively. He is currently working toward the Ph.D. degree at Delhi Technological University, New Delhi, India. He has published/presented few technical papers and conference proceedings in international journal of repute. His research interests include electric vehicles and optimization techniques.

M. Rizwan did his post-doctoral research at Virginia Polytechnic Institute and State University, USA. He has more than 19 years of teaching and research experience. Dr. Rizwan has successfully completed three research projects in the area of renewable energy systems and published and presented more than 155 research papers in reputed international/national journals including IEEE transactions and conference proceedings. He is the recipient of Raman Fellowships for Post-Doctoral Research for Indian Scholars in USA, DST Start Up Grants (Young Scientists) and many more. His area of interest includes power system engineering, renewable energy systems particularly solar photovoltaic, building energy management, smart grid and soft computing applications in renewable energy systems. He is Sr. Member of IEEE, Life Member, ISTE, Life Member SSI, Member International Association of Engineers (IAENG), and many other reputed societies.

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

Intelligent Chatbot Using GNMT, SEQ-2-SEQ Techniques

Rajesh Kumar Yadav
Department of CSE
Delhi Technological University
New Delhi, India
rkyadav@dtu.ac.in

Sahil Harwani
Department of CSE
Delhi Technological University
New Delhi, India
sahilharwani1999@gmail.com

Satyendra Kumar Maurya
Department of CSE
Delhi Technological University
New Delhi, India
mauryasats98@gmail.com

Sachin Kumar
Department of CSE
Delhi Technological University
New Delhi, India
sachink061198@gmail.com

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498485

***Abstract*—Dialogue generation and intelligent chatbot development is a fairly new avenue in the field of artificial intelligence and machine learning. Chatbot development is a popular field in natural language processing. Nowadays a large amount of research projects are currently underway to make efficient and reliable chatbots using machine learning, artificial intelligence and natural language processing. Chatbots or dialogue generation agents are mainly used by businesses like online stores or startups, financial institutions like banks, credit card companies and govt organisations. There are many different types of chatbots available differing in frameworks such as code based or interface based. But these chatbots lack in genuine dialogue generation and seem rather bland and unintuitive. Some of the popular intelligent chatbots or you can say personal assistants include google assistant, Siri by apple and cortana by microsoft. The functioning of these chatbots is constrained and limited and they are not aimed to hold real world conversations. In this project endeavour we have tried to develop an intelligent chatbot using state of the art GNMT(Google neural machine Translation) model which is built upon seq-to-seq modelling having encoder- decoder architecture. This encoder decoder discussed here uses recurrent neural networks along with bi directional LSTM(Long short term memory) which is one of the hot areas.**

***Index Terms*—GNMT, seq-to-seq, LSTM**

I. Introduction

In today's rapidly growing technologies, artificial intelligence and machine learning are two of the fastest growing fields. One of the recent advancements include "chatbot" or "bot". These chatbots are used to engage with potential clients or customers on the business websites. We can use these chatbots to attain some of the very useful information about the client, like what kind of product he is interested in, how much quantity he wants to buy, what are his buying capacities and much more insightful information to make the perfect deal with the customer. Chatbots are also widely used on social media platforms like Instagram, Facebook and LinkedIn etc. They rely on the similarity of keywords and retrieved texts from external or internal data sources, which also includes the organizational database of the World Wide Web. With natural language processing techniques and machine learning (ML) algorithms, some other advanced chatbots have been created. There are also many commercial chat engines available that help create chatbots based on client databases. Their operation can vary from customer service to a personal assistant, product suggestion, product inquiry. Many of these chat agents are built using rule-based methods, recovery techniques or simple algorithms for machine learning. They can be used to make connections with other people by chatting with them and thereby making a connection with them. In this project, we are going to do some extensive research on the review papers published on "sequence to sequence" technique and "google's neural translation machine". Our research will be centered around these two technologies and we hope to find some novel technique of developing chatbots. Since chatbots imitate an individual human, they are designed using Artificial Intelligence (AI) techniques. Deep learning that imitates the human brain itself is a similar technique within AI. It detects training patterns in data and uses the same features to analyze new data.

II. Related Work

For the better understanding of the concepts that we used in our chatbot development, we have done extensive research in the topics related to chatbot development. The initial methods of chatbot development included the use of standard STM-based systems for dialogue generation. The major reason behind this was the unlimited memory capacity of STM-based systems for a large dataset. Recent advances in the field show that a simple LSTM approach gave more accurate results. It was proposed that the use of LSTM approach would be significantly better if the model could be trained on a big data-set.[2] Another interesting module that is the machine translation module, also came up recently. Although the module was originally built for the purpose of translating sentences from one language to another, it has great potential for dialogue generation as well.[3] Many recent advances and experiments in conversational agents have been carried out in the Framework. In addition to standard chatbot creation techniques which use rule-based techniques or basic machine

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:55:23 UTC from IEEE Xplore. Restrictions apply.

learning algorithms, a lot of advanced chatbots also us advanced natural language processing techniques and dee learning techniques, such as advanced natural language pro cessing (NLP) techniques, Learning with Deep Reinforcemer (DRL) and Deep Neural Network (DNN).

III. Problem Statement

We have come to the conclusion with the help of our surve paper that a combination of google's neural machine transl tion and sequence to sequence will result in a chatbot wit better accuracy and better sentence development. Now th we have produced has led us to our problem statement. Now since we have an authentic problem statement and resources to go for a solution. The problem statement helped us to be clear with our prerequisites, challenges and direction. Earlier chatbots have been developed with deep learning technology giving not much accuracy while translation which is BLEU score.

IV. Methodology

A. Sequence to Sequence (Seq2Seq)

An encoder-decoder based model is such a structure that is majorly used in generating dialogue, machine translation and language modeling. Seq-2-Seq uses Recurrent Neural Network mainly for Natural Language Processing activities, which is a standard form of the Deep Neural Network. More or more RNN types are used for decoders in the sequence to sequence model (Seq2Seq). The input sequence is inserted as the text vector indicator in the encoder in this encoder-decoder construction. The coder then displays an intermediate representation of the vectors of information or reasoning. As a result, this thought vector produced by the code is saturated as input to the decoder. Finally, in the form of an objective sequence, the decoder analyzes the imaging vector and converts a single-term sequence and then generates several outputs from the decoder. While vanilla RNN is automated in Seq2Seq and it works very well on most of the NLP problems to date, vanilla network cells often fail due to the extremely high language modelling issue, particularly when long-term sequences need to be retained, as the information is always remembered as large in large data sets and then turned into an RNN network information bottle. Therefore, researchers are using a variety of repeated neural networks to address such problems.

B. Deep Reinforcement Learning

Deep Reinforcement Learning has been extensively used by Dan Jurafsky for Dialogue Generation to create lengthy communication conversations. The Seq2Seq model can generate consistent conversations, but it can produce standard repetitive responses irrespective of input and also can be caught in long conversations being in loop. This happens when one Seq2Seq simultaneously predicts speaking while ignoring their potential for future performance. Seq2Seq versions seem to produce repetitive responses such as; which are very common. This has happened due to the high frequency of standard responses

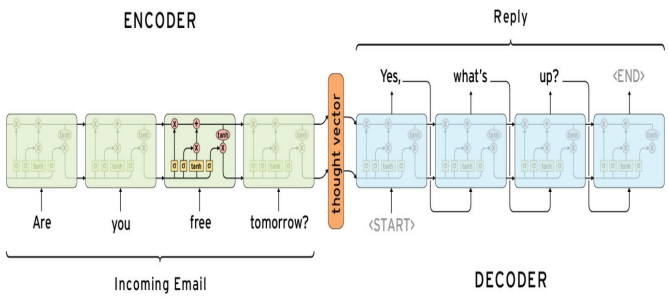


Fig. 1.

in the training set, and these responses are highly consistent with a large range of input text. We started with reading some papers regarding the seq-to-seq module, GNMT module and chatbot developed using them. That gave us the motivation to explore more in this field. Later we finalised three research papers and read about the implementation, prerequisites, scope of research in this area.

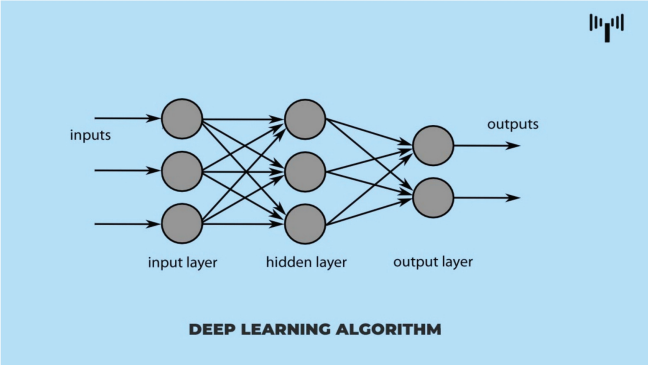


Fig. 2.

C. Google's Neural Machine Translation(GNMT)

GNMT is a module that is used for neural machine translation into a variety of languages including English. GNMT has also been used to create dialogues. It is based on the popular seq2seq technique. GNMT contains an array of modules which can intelligently create dialogues. Modelling seq to seq with an encoder-decoder framework designed using Long short-term memory requires GNMT. In terms of their exact architecture, NMT models differ. The RNN, utilised by frequent NMT models, is a natural option for sequential data.

For both encoder and the decoder, a Recurrent Neural Network is usually used. Moreover, the Recurrent Neural Network differ in features of: (a) bidirectional or directionality-unidirectional; (b)multi-layer or depth-single; and (c) type-often either a Long Short-term Memory (LSTM), vanilla RNN or a gated recurrent unit (GRU).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:55:23 UTC from IEEE Xplore. Restrictions apply.

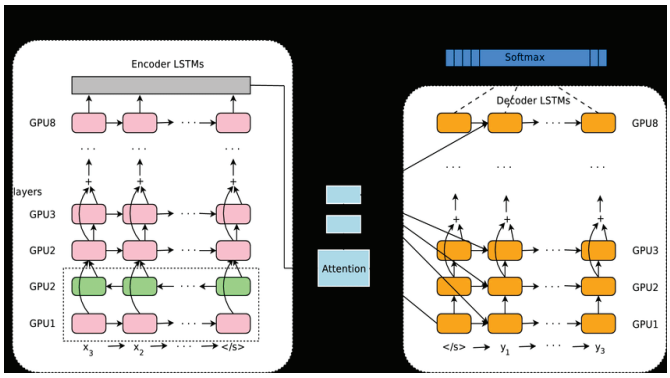


Fig. 3.

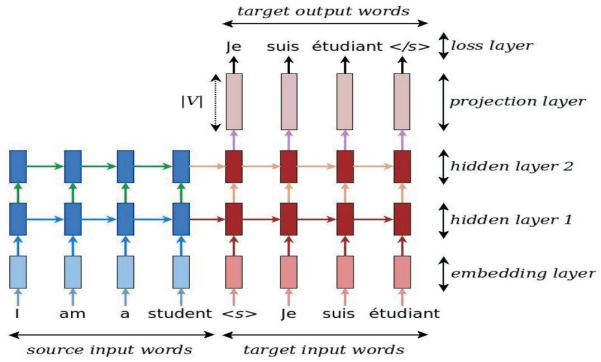


Fig. 5.

D. Neural Attention Mechanism

Constructing modern machine translation systems, attention-grabbing works perfectly well, it was introduced in 2015. The concept of how to pay attention creating a short direct link between the source and the target by paying "Attention" to the corresponding source content while it translate .

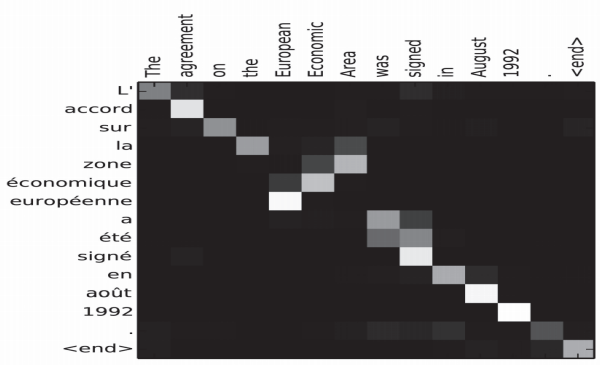


Fig. 4.

Attentional view - an instance of the correlation among source and purpose sentences. In the model vanilla sequence to sequence, the final state(source) from encoder to decoder is past when the translation begins. This functions fine for medium and short lengths text strings. Moreover, in big sentences, one four-dimensional state - the hidden xed-size becomes bottle information. Instead of removing all concealed regions included in RNN source, the attention method gives a method that permits the decoder to see it (treat it as a powerful memory of source information). This is the mechanism of attention the way it works enriches the conversion of long sentences. Here are some suggestions on how to look or get an appointment for acne treatment sector-level and applied to other NLP activities including photo captioning, speech recognition, and summary of the text.

Attention method - an example of an NMT based program.

1. The current concealed state is measured against all resource sources for attention. 2. Based on the weight of attention we calculate the vector of the content as the rated rating the source says. 3. Merge the context vector and the hidden present condition of the initial output attention vector.

E. Long short-term memory

Long Short-Term Memory is a special type of Recurrent Neural Network(RNN) cell, namely resolves the bottle problem in a long sequence. LSTM forgot gate and installation and exit gates. This makes it easier to recall long series without overburdening the network by removing irrelevant information.

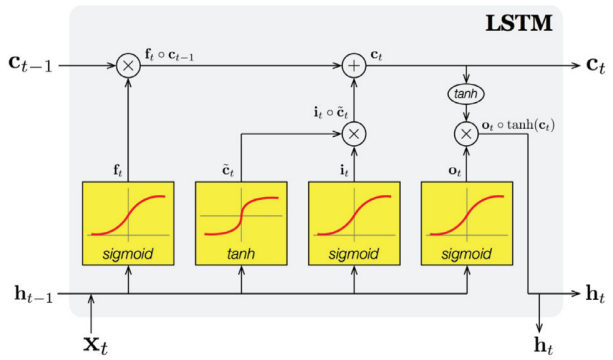


Fig. 6.

(1) $f_t = \sigma_g(W_f\ x_t + U_f\ h_{t-1} + b_f\ )$
(2) $i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i)$
(3)$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o)$
(4) $c_t = f_t.c_{t-1} + i_t.\sigma_g(\sigma_g(W_c x_t + U_c h_{t-1} + b_c)$
(5) $h_t = o_t.\sigma_h(c_t)$

Variables

$x_t \in R^d$: input vector to the LSTM unit $f_t \in R^h$: forget gate's activation vector $i_t \in R^h$: input gate's activation vector $o_t \in R^h$: output gate's activation vector
$h_t \in R^h$: output vector of the LSTM unit
$c_t \in R^h$: cell state vector
$W\ R^{h\times d}$, $U\ R^{h\times h}$ *and* $b\ R^h$:matrix weight and bias vector parameter which needs to be read during training when high-d

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:55:23 UTC from IEEE Xplore. Restrictions apply.

d and h refers to the number of input features and number of hidden units, respectively.

V. Implementation

- In this project following the state of the art strategies proposed in the recently published research papers, we used the Neural Machine Translation (NMT) module. Although the Neural Machine Translation (NMT) module module is used for machine translation tasks, other NLP functions, including dialogue and text summary, have been effective. In chat production, Neural Machine Translation (NMT) module has a rich Seq2Seq module with many additional features. In conducting dialogue and multitasking NLP activities, Seq2Seq is an industry standard. TheGoogle's Neural Machine Translation (GNMT) module model incorporates the Sequence to Sequence sequence (Seq2Seq) based on uni cells or the LSTM side with encoder-decoder structures.
- Data Collection and Preprocessing For this project, We have used "Cornell Movie Subtitle Corpus" database for training in the final model. Preprocessing of Cornell Movie Subtitle Corpus:- Movie ID, Character ID, and Movie Line ID is divided in the conversation details in the movie corpus by + +++++ ". Going forward, chat data is cleared to reject this meta-data (e.g. movie ID, character ID, queue ID). Alongside, data separators ("+++++) have also been removed. In addition, another character in the data contains a coding format that is not supported by the UTF-8 standard and was later removed from there. Ultimately, the data was split into two different files to fit Google's format Neural Machine Translation (GNMT) machine input format where first file is the dialogue 1 and second format is to respond to the dialogue 1. Following the separation of the two archives, the data in both files was cleared at the same time.. Everything that is out of order like, out of alphabetical order, other punctuation marks (.,?! ') have been erased as they were irrelevant to the conversation. In fact the whole text has been changed to lowercase . After that, multiple the subsequent occurrence of these symbols (.,?! ') is reduced to one for reduction fullness of punctuation marks. After that entire punctuation marks (') are separated by only a single space in the GNMT module before and after better performance. Finally, everything is in order: the space is reduced to one space and the string of each text is cut to remove the front and back of the space.
- Also, data is cleared by deleting external conversations. If there are multiple speeches in a row. From one person there was everything except the last person 's speech. Initially, speeches over 100 were discarded in both text conversations and their response as the text length increases, contextual consistency begins to decline due to variability and limited data. But in time, the full length of the text was embedded. After clarification, the source and the intended concept were divided into training, testing and development / verification set by source and target format and saved to files' final input pipe input. As suggested by Google tensorflow , Google's Sub-word Neural Machine Translation (NMT)is used for word processing and GNMT module texts.
- In the second phase we made our model which is a combination of both nmt and seq-2-seq.
- In the third phase, we trained our model based on the data collected. We used 30 percent of data for training and others for evaluation.
- In the fourth phase testing and response detection was done. Model testing was done using the given input and the expected out. A detailed input-expected chart was prepared for this.

VI. Result Analysis

Following are the responses received after the training to a complete database with 34MB of training text containing 225000 + 225000. The first test result produced moderate combination sentences. The following responses were made after the discovery by a trained model. Assuming, a trained model has produced a response of 10-30 for each input. Therefore, the following examples are best suited from 10-30 students' responses during consideration.

TABLE I

Input (Person 1)	GNMT Model Output (Person 2 Response)
how are you	how am i?
whats your plan for the weekend ?	i dont know
would you like some tea	would you like some help ?
You got something on your mind	yeah
would you like to go to watch movie ?	no .
where would you like to go for lunch today	i know where to go for lunch today .
i know where to go for lunch today .	you know where to go ?
Where is the library located ?	is that where you were ?
weather seem warm would you like to get some cold drinks	would you like to get some cold coffee ?
did you like the new startrek movie sequel	i did .
band played really well i am going for their next tour	i am going for their next interview .
am i disturbing you ?	no . not at all . where are you ?
i'm heading out . how about you ?	i got to wait for a call .

Next up is Perplexity and Bleu for testing and development data.

TABLE II

Dataset	Perplexity	Bleu
eval dev	50.76	10.1
eval test	46.82	10.6

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:55:23 UTC from IEEE Xplore. Restrictions apply.

VII. Conclusion

In creating chatbots or chat generators, the challenge lies in building compatible systems for coherent dialogue generation. Because machine translation is the model used in this test, the production of dialogue is considered a translation issue, where the history of previous conversations is ignored. Thus in relation to long conversations, the model can be limited in functionality. With real-life high-quality conversational datasets that can better mimic human interactions, a chatbot created using Google's Neural Machine Translation Model (GNMT) can be further enhanced. The GNMT model's hyper-parameters can also be optimized and optimized for performance improvement. A combination of GNMT along with seq2seq could be used depending on the opportunities available to advance the project, which would significantly improve the chat bot's performance. This was an attempt to hypothesize that GNMT would prove useful for dialog generation in order to develop intelligent chatbot along with seq2seq methods.

References

[1] Li J, Monroe W, Ritter A, Galley M, Gao J, Jurafsky D. *Deep reinforcement learning for dialogue generation.* arXiv preprint arXiv:1606.01541. 2016 Jun 5.

[2] *Google's Neural machine translation(NMT)* https://github.com/tensorflow/nmt

[3] *Neural Machine Translation (seq2seq) Tutorial* https://www.tensorflow.org/ tutorials/seq2seq

[4] *seq2seq Github Repository* https://github.com/google/seq2seq

[5] *Google's seq2seq* https://google.github.io/seq2seq/

[6] *Cornell Movie Dialogs Corpus* http://www.cs.cornell.edu/∼cristian/Cornell_Movie-Dialogs_Corpus.html

[7] https://medium.com/swlh/chatbots-of-the-future-86b5bf762bb4

[8] https://chatbotsmagazine.com/to-build-a-successful-chatbot-ask-these-5-questions-b7fe3776c74c

[9] https://labs.bawi.io/creating-a-conversational-chatbot-using-wit-ai-6eba3c625f4f

[10] https://www.ipsoft.com/2017/11/20/when-chatbots-fail-virtual-agents-step-in/

[11] https://isaacchanghau.github.io/2017/08/02/Seq2Seq-Learning-and-Neural- Conversational-Model/

[12] https://chatbot.fail/

[13] https://www.altoros.com/blog/text-prediction-with-tensorflow-and-long-short-term-memory-in-six-steps/

[14] https://medium.com/the-mission/11-best-uses-of-chatbots-right-now-1c27764b7e62

[15] https://www.marutitech.com/7-reasons-why-business-needs-chatbot/

[16] https://chatbotsmagazine.com/how-to-develop-a-chatbot-from-scratch-62bed1adab8c

[17] https://www.marutitech.com/complete-guide-bot-frameworks/

[18] https://en.wikipedia.org/wiki/Long_short-term_memory

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:55:23 UTC from IEEE Xplore. Restrictions apply.

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

Check for updates

JOURNAL OF ENERGY RESOURCES TECHNOLOGY

Investigation of a combined refrigeration and air conditioning system based on two-phase ejector driven by exhaust gases of natural gas fueled homogeneous charge compression ignition engine

Abdul Khaliq
Department of Mechanical Engineering, College of Engineering at Yanbu, Taibah University, Yanbu Al Bahr 41911, Saudi Arabia
Email: khaliqsb@gmail.com

Bandar A. Almohammadi
Department of Mechanical Engineering, College of Engineering at Yanbu, Taibah University, Yanbu Al Bahr 41911, Saudi Arabia
Email: bmohamdi@taibahu.edu.sa

Mathkar A. Alharthi
Department of Chemical Engineering, College of Engineering at Yanbu, Taibah University, Yanbu Al-Bahr 41911, Saudi Arabia, Email: mathkar.alharthi@gmail.com

Mohd Asjad Siddiqui

Department of Mechanical Engineering, Delhi Technological University, Shahbad Daulatpur, Bawana Road, Delhi-110042, India

Email: siddiquimohd.asjad@gmail.com

Rajesh Kumar[1]
Department of Mechanical Engineering, Delhi Technological University, Shahbad Daulatpur, Bawana Road, Delhi-110042, India
Email: dr.rajeshmits@gmail.com

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Abstract

A natural gas fueled homogeneous charge compression ignition (HCCI) engine is coupled to an exhaust gas operated turbine driven two-phase ejector cycle to generate power and cooling energy, simultaneously. By establishing a thermodynamic model, the simulation of the proposed system and its parametric analyses are conducted. Energetic and exergetic investigations are carried out to study the role of equivalence ratio, engine speed, condenser temperature, refrigeration evaporator temperature, air conditioning evaporator temperature, and ejector nozzle efficiency on the thermodynamic performance parameters of the combined cycle. The analysis of two-phase ejector cooling cycle using three working fluids including R717, R290, and R600a is conducted. Results reveal that the thermal efficiency of HCCI engine is increased from 47.44% to 49.94%, and for the R600a operated combined cycle it is increased from 60.05% to 63.26% when the equivalence ratio is promoted from 0.3 to 0.6. Distribution of fuel exergy results show that out of 100% exergy input, in case of R717 operated combined cycle, 139.79 kW (38.72%) is the total exergy output and 164.21 kW (45.49%) and 57 kW (15.79%), are the values for exergy destruction and exergy losses. It is further shown that change in refrigerant minorly influence the percentages of exergy distribution.

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

***Keywords**: HCCI engine, two-phase ejector cycle, refrigeration, air-conditioning, natural gas*

[1]Corresponding author Email: dr.rajeshmits@gmail.com (Rajesh Kumar)

1. Introduction

The increasing rate of the global energy consumption for refrigeration and air conditioning is a basic issue that our society has to manage due to the consequent problems, as the greater increase of greenhouse gas emissions and fast depletion of fossil fuel reserves [1]. Commitment to meet out the refrigeration and air conditioning needs in a sustainable manner is a climate change challenge. In this regard, waste energy recovery from prime movers such as; gas turbine, internal combustion engines, and steam power plants creates an appealing opportunity of energy saving and mitigation of carbon footprint [2]. Moreover, development of cooling-power cogeneration system provides an assistance in resolving the concern related to energy crisis and climate change. Producing the refrigeration is more expensive than heating and power generation in majority of applications as it requires refrigeration machine and electricity. Therefore, recovery of waste heat from the engines for the additional generation of power and cooling energy is capable to provide economic benefits [3,4]

Maintenance of quality through preservation of food products which are transported through road vehicles with comfortable journey requires cooling for refrigeration and air conditioning. In the land transportation vehicles such as diesel engine aboard trucks, approximately two-thirds of the total energy input is lost as heat rejection to the environment because of the high exhaust gas temperature which usually varies between 400^0C and 600^0C [5,6]. Therefore, development of systems that can convert the engine waste heat directly into cooling in various range of temperatures provide a promising means of effectively meeting the cooling demand of land transportation vehicles. Cooling systems powered by waste heat of internal combustion engines

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

have been widely investigated by many researchers in the recent past. Cooling systems which are often used as bottoming cycles in the internal combustion engine are based on absorption refrigeration which is considered as a cooling technology where the working fluid pair (strong solution) at the expense of thermal energy of the heat source get separated into the absorbent and refrigerant. Finally, the refrigerant takes heat to refrigerate the required thermal load at the evaporator. Most commonly chosen working fluid pairs for absorption refrigeration systems (ARS) are LiBr-H_2O and NH_3-H_2O that can provide cooling in different range of temperatures. In the LiBr-H_2O operated ARS, use of water as the refrigerant makes the system un-functional when the evaporator temperature is lower than 0^oC. For producing the refrigeration below 0^oC, NH_3-H_2O operated ARS is commonly used, and this in general, is the case of industrial refrigeration [7]. However, in the power cycles or internal combustion engines combined with NH_3-H_2O operated ARS, a rectifying column is required which results in the reduction of system performance. Moreover, when the high temperature engine exhaust gas is used to drive the absorption refrigeration cycles directly, the serious temperature mismatch between the heat source and the working fluid results in huge exergy destruction [8]. Further, ARS are largely used for exhaust heat recovery applications in industrial and marine applications. It is scarce to have a waste heat recovery from automobile engines using ARS.

The limitations accompanied by absorption refrigeration systems created a need for the search of alternative energy-efficient and sustainable cooling solutions. In this regard, ejector refrigeration cycle (ERC) has recently been emerges as one of the very attractive cooling technology options because it offers better energy efficiency and low cost of installation. Moreover, its operation requires no mechanical power and shows a flexibility in operation with various kind of refrigerants [9]. In addition to this, ejectors have been found as compact and easier in maintenance compared

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

to traditional cooling systems [10-14]. In the literature, many studies are reported regarding the application of ejector cooling technology for achieving the processes of sustainable refrigeration. The investigations devoted to the study of ERC driven by the exhaust gases of internal combustion engine are relatively rare. Yilmaz and Aktas [15] investigated the performance of an ejector refrigeration (ER) system driven by the exhaust waste heat of a heavy vehicle engine. In their configuration, refrigerants R134a and R245fa were used for the comparative simulation and their results show that, the performance of the system would be higher if R245fa is preferred rather than R134a with the given operating conditions. Jaruwongwittaya and Chen [16] investigated the feasibility of implementing a two stage ejector cooling system in a bus using water as a working fluid. Their results indicated that the application of two-stage ejector cooling system to replace the vapor compression system for engine waste heat recovery is a suitable alternative and its applications leads to the reduction of fuel consumption, green-house gas reduction as well as resolve the problem of a high condensing temperature of the cooling system in a bus engine. Xia et al. [17] developed a combined power and cooling system for engine waste heat recovery which comprises an organic Rankine cycle (ORC) and an ejector refrigeration cycle (ERC). The system overall performance was evaluated and an optimization was carried out to search the optimal system performance from exergy point of view. Two-phase ejector refrigeration systems have been largely investigated theoretically and experimentally, and in this regard, following studies are found. Kornhauser [18] presented a study conducted on different working fluids in vapor compression refrigeration system using two-phase ejector. Chaiwongsa and Wongwises [19] investigated the performance of ejector refrigeration cycle which consists of a single evaporator and a refrigerant separator. The effects of exit diameter of nozzle and the throat section of the ejector was ascertained. Pottker et al. [20] performed an experimental investigation on ejector

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

cooling system and carried out the analysis without considering the ejector dimensions. Their analysis reveals that application of R410A refrigerant in the system provide great thermodynamic benefits. A two-phase ejector refrigeration cycle operated on low pressure refrigerants was investigated theoretically and experimentally by Lawrence [21]. A theoretical comparative study on the use of R143a and R1234yf as refrigerants in ejector cooling system was conducted by Lawrence and Elbel [22]. An increase in the COP of 5% and 6%, respectively, was found. Wang and Yu [23] examined the effect of component efficiencies on the performance of two-phase ejector. In their investigation, a new method of calculation for the component efficiencies of ejector was proposed by combining ejector model with the reported data. Zengenhagen and Ziegler [24] presented the assessment of an engine exhaust heat driven jet-ejector cooling system integrated to turbocharged gasoline engines. They determined the energy efficiency and charge air temperature for the use of R134a as the refrigerant in ejector cooling. Unal [25] developed the exhaust heat driven two-phase ejector cooling system for reducing the load and fuel consumption of the engine of passenger bus. The author stated that application of two-phase ejector as an expansion valve can enhance the performance of bus air conditioning system. Zhang et al. [26] developed an exhaust heat recovery unit for the automobile integrating to ERC. They analyzed the performance of ejector cooling system and recommended a suitable range of operating parameters for the efficient operation of combined system. Above reported investigations proved that two-phase ejector cooling systems has a great potential for the recovery of engine exhaust heat. Siddiqui et al. [27] proposed a novel combined power and cooling system consisting of the ejector refrigeration cycle (ERC) and organic Rankine cycle (ORC) to enhance the overall efficiency of wet-ethanol fueled homogeneous charge compression ignition (HCCI) engine. Their results show that increase in evaporator pressure of ERC is of the great benefit of thermodynamic performance

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

of cogeneration system because its elevation significantly increases the cooling capacity. Yu et al. [28] presented a novel cooling-power cogeneration system consisting of the ORC and ERC for the efficient utilization of low grade waste heat. They optimized the system performance and working selection for the fixed cooling outputs. Their investigations found perfluoropropane as the most suitable working fluid for the proposed system.

Studies carried out on the recovery of thermal exhaust of alternative fueled internal combustion engine reveals that integration of two-phase ejector with the natural gas fueled engine for the simultaneous production refrigeration and air conditioning (for food preservation and cabin cooling) has not to date been reported in the literature. Moreover, proposal and analysis of a two-phase ejector cooling system applied for exhaust heat recovery from the relatively new combustion technology engine that runs on lean air-fuel mixture and produces least harmful emissions while maintaining the higher engine efficiency is very much missing in the literature. Therefore, employment of a shaft power driven two-phase ejector which consists of a hermetic reciprocating compressor, an air cooled condenser, a separator, and two evaporators for combined production of refrigeration and air conditioning and integrated to HCCI engine operated on natural gas has been found worth and novel for current investigation. The development of such a cooling-power cogeneration system satisfy the diverse consumer's demands and is having three fold benefits; cooling produced by low temperature evaporator will refrigerate a thermal load for food and vaccine preservations and the cooling produced by high temperature evaporator will provide cabin cooling or air conditioning of vehicle, and decrease in exhaust temperature due to waste heat recovery will reduce the thermal pollution. Proposed system is investigated parametrically to study the role of key operating variables such as; equivalence ratio, engine speed, refrigerant condenser temperature, efficiency of ejector nozzle, temperature of refrigeration evaporator, and air

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

conditioning evaporator on the thermal and exergetic efficiencies of combined cycle as well as on the efficiencies of natural gas fueled HCCI engine.

2. Proposed combined cycle description

Schematic diagram of the combined cycle proposed in this study is displayed in Fig. 1. The proposed system comprised of a natural gas-fueled HCCI engine and a two-phase ejector system and the combined cycle generates power and cooling for air-conditioning and refrigeration, simultaneously. The main components of the proposed system are as follows: two compressors, regenerator, fuel vaporizer, HCCI engine, catalytic convertor, turbine, heat exchanger, an air-cooled condenser, an air-conditioner, an ejector, a refrigerator, a separator, and two expansion valves.

Ambient air (1) enters the compressor (C1) where it is compressed to higher pressure (2) and the passes through regenerator (Reg) which increases its temperature (3) and then mixed with natural gas (4) injected into the fuel mixer (FM). In the fuel mixer, compressed air and natural gas mixture forms a homogeneous air-gas mixture (5) and enters to the HCCI engine. This homogeneous mixture and residual gases mixed in the cylinder and engine exhaust gases enter (6) to the catalytic convertor (CC). The unburned hydrocarbon (UHC) and carbon monoxide (CO) combust in the convertor (CC) which were not burned in the combustion chamber and increases exhaust gas temperature due to heat release. The less harmful gases from the catalytic convertor (CC) enter (7) to the turbine (T) and generates power to drive the compressors (C1 & C2). The exhaust gases from the turbine (T) enter (8) the regenerator (Reg) and exchange heat with the compressed air coming from the compressor (C1) and leaves the system (9).

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

The compressed air from the compressor (C2) enters the condenser (Cond) (10) where heat is rejected out to the environment. The refrigerant leaves the condenser as saturated liquid at (11) and goes to the air- conditioner (AC), where it evaporates completely to vapor for air-conditioning (13) through the heat exchanger (HX) (11-12) and expansion valve (EV1) (12-13). A heat exchanger (HX) between the expansion valve (EV1) and the condenser (Cond) transfers heat from the saturated liquid leaving the condenser (Cond) to the saturated vapor going into the compressor (C2). The two-phase fluid at the air-conditioner (AC) exit (14) flow into the ejector (Eje) motive nozzle. The high-velocity vapor (14a) entrains secondary fluid from the refrigerator (ref) (18) and mixes with the primary fluid. Stream (14a) and (18a) mixed together in the mixing chamber, resulting in a homogeneous mixture. Pressure increases due to normal shock when mixed fluid enters the constant area section. The mixed fluid (15m) decelerates in the diffuser after the shock. The mixed two-phase fluid from the ejector (Eje) (15) enters the separator (Sep). The saturated liquid fluid at the bottom of the separator (Sep) (16) flows into the refrigerator (ref) to produce desired refrigeration effect (17) through the expansion valve (EV2) and the saturated vapor fluid (19) is heated when it goes through the heat exchanger (HX) and then enters to the compressor (C2) (20).

3. Thermodynamic properties evaluation

In the internal combustion engines operated on ideal cycle, air is considered as working fluid for the processes occurring in the engine cylinder. However, the working fluid appears to be a mixture of many gaseous species in the actual engine cycle operation.

Variation of specific heat of gases with temperature is quite considered in the actual cycle operation by ignoring its variation with pressure. The molar specific heat model considered in the

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

current study consisting of the variation of $\bar{C}_p(T)$ with temperature in the range 300K to 3000K are available in the Ref. [29]

$$\bar{C}_p(T) = \bar{R}(a_1 + a_2T + a_3T^2 + a_4T^3 + a_5T^5) \quad (1)$$

The relation of specific heat with two most prominent thermodynamic properties such as enthalpy and entropy can be found as [30]

$$\bar{h}_{i_{(T)}} = \bar{h}_{i_{(T_0)}} + \int_{T_0}^{T} \bar{C}_{p_i}(T)dT \quad (2)$$

$$\bar{s}_{i\ (T,p)} = \bar{s}_{i_{(T_0,p_0)}} + \int_{T_0}^{T} \frac{\bar{C}_{p_i}(T)}{T} dT - \bar{R}ln\left(\frac{y_i p}{p_0}\right) \quad (3)$$

In the current investigation, flow streams such as exhaust gases, products of combustion, and the air inducted are assumed as gaseous species mixture. Commonly used thermodynamic properties of the mixture can be expressed as [30].

$$\bar{h}_{mix_{(T)}} = \sum_{i=1}^{n} y_i \bar{h}_{i_{(T)}}; \qquad \bar{s}_{mix_{(T,P)}} = \sum_{i=1}^{n} y_i \bar{s}_{i_{(T,P)}}; \qquad M_{mix} = \sum_{i=1}^{n} y_i M_i \quad (4)$$

The combustion species produce under ideal condition are; water (H_2O), carbon di oxide (CO_2), and nitrogen (N_2) from the air. But for the actual combustion process, engine exhaust gases contain several other compounds along with H_2O, CO_2 and N_2. The combustion of natural gas in the HCCI engine is assumed to produce the following combustion products as: CO, H_2O, CO_2, O_2, H_2, and N_2. After considering methane as the main component of natural gas, its combustion with air results in the formulation of the equation which can be expressed as [31]

$$CH_4 + \left(\frac{2}{\emptyset}\right)(O_2 + 3.76N_2) \rightarrow n_1CO_2 + n_2H_2O + n_3N_2 + n_4O_2 + n_5CO + n_6H_2 \quad (5)$$

The mole fractions, n_i, combustion product species have been computed under the conditions of equilibrium.

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

Emissions from the HCCI engines such as carbon monoxide and unburned hydrocarbons can undergo the process of further oxidation by using Pd-loaded $SiO_2 - Al_2O_3$ as the oxidizing catalyst. By employing the equations, the composition and temperature of outlet gases can be calculated as [32]

$$\sum_{Product} n_e\left(\bar{h}_f^{\circ} + \Delta\bar{h}\right)_e = \sum_{Reactant} n_i\left(\bar{h}_f^{\circ} + \Delta\bar{h}\right)_i \tag{6}$$

3.1 Turbocharger compressor. After considering the fixed value of polytropic efficiency at a given pressure, the air temperature at the outlet of compressor can be calculated after employing the equation [33],

$$\mathrm{T}_2 = \mathrm{T}_1\left(\frac{\mathrm{P}_2}{\mathrm{P}_1}\right)^{\frac{(\gamma-1)}{\gamma\eta_C}} \tag{7}$$

Energy balance applies to compressor gives

$$\mathrm{w}_c = \mathrm{h}_2 - \mathrm{h}_1 \tag{8}$$

3.2 Regenerator. The compressed air is preheated in the regenerator whose performance is evaluated using its effectiveness and it can be shown as

$$\varepsilon_h = \frac{\left[\left(\dot{m}.\mathrm{c}_p\right)_2(\mathrm{T}_3-\mathrm{T}_2)\right]}{\left[\left(\dot{m}.\mathrm{c}_p\right)_{min}(\mathrm{T}_8-\mathrm{T}_2)\right]} \tag{9}$$

The assumption, $\left(\dot{m}.\mathrm{c}_p\right)_2 = \left(\dot{m}.\mathrm{c}_p\right)_{min}$ is considered to calculate the temperature, T_3, the temperature of air at the outlet of regenerator is determined after applying the equations given below

$$\mathrm{T}_3 = \mathrm{T}_8\varepsilon_h + \mathrm{T}_2(1-\varepsilon_h) \tag{10}$$

$$\mathrm{H}_3 - \mathrm{H}_2 = \mathrm{H}_8 - \mathrm{H}_9 \rightarrow \mathrm{H}_9 = \mathrm{H}_8 - \mathrm{H}_3 + \mathrm{H}_2 \tag{11}$$

3.3 Fuel-air mixer (FM). The temperature at the outlet of fuel-air is determined after applying

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

$$H_5 = H_3 + H_4 \tag{12}$$

3.4 HCCI engine. The gaseous mixture from fuel-air mixer is delivered to HCCI engine and its thermodynamic cycle is given in Fig. 2. The relation applied to compute the temperature of the mixture for intake process $(i - 1')$ can be expressed as [31].

$$T_{1'} = \frac{T_i(1-f)}{1-\frac{1}{(n.r_c)\left[\frac{p_e}{p_i}+(n-1)\right]}} \tag{13}$$

Where n is the average specific heat capacity ratio which was taken as 1.35 for lean combustion, f was assumed as 0.03 and p_i/p_e = 1.4 [34].

The pressure remains unchanged:

$$P_{1'} = P_i \tag{14}$$

The mixture following the process of compression $(1' - 2')$ and combustion, and properties at 2' can be determined.

Properties after completing the process of heat addition (2'-3') are computed by

$$\dot{Q}_{in} = \dot{Q}_{fuel} - \dot{Q}_l \tag{15}$$

Where $\dot{Q}_{fuel}$ is given as

$$\dot{Q}_{fuel} = \eta_{comb}\dot{m}_{fuel}Q_{LHV} \tag{16}$$

The losses inside the combustion chamber are evaluated by considering the combustion efficiency [35]

$$\eta_{cc} = 100 - V_{\%} - (1.26 + 0.25C_{e2} + 0.4C_{e2}^2) + (\emptyset - 0.2486)8.1 \tag{17}$$

Where $C_{e2} = Ln(V_{\%})/Ln(2)$

$\dot{Q}_l$ (heat losses from cylinder walls) is determined using:

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

$$\dot{Q}_l = h_c * 10^{-3}(A_{ch} + A_{cyl})(T_{avr} - T_w)/(2 * \dot{m}_a) \quad (18)$$

Where, h_c is the heat transfer coefficient and given as [36].

$$h_c = (3.26L^{-0.2}p^{0.8}\mathrm{T}_{avr}^{-0.73}\omega^{0.8}) \quad (19)$$

Where,$\omega = 2.28\bar{s}_p + (3.24 * 10^{-3}/6)(V_d T_{ref}(p - p_{mot})/(p_{ref}V_{ref}))$ (20)

Where p_{mot} is the pressure at motored conditions [37]:

$$p_{mot} = p_{ref}(V_{ref}/V)^n \quad (21)$$

The parameters from exit of expansion process (3'-4') are determined similar to the method adopted for the compression stroke (1'-2').

The parameters at the outlet of blowdown process are computed using:

$$p_{5'} = p_e \quad (22)$$

$$T_{5'} = T_{4'}(p_{4'}/p_e)^{(1-n)/n} \quad (23)$$

$$v_{5'} = \frac{\bar{R}}{M_{mix}} * \frac{T_{5'}}{p_{5'}} \quad (24)$$

Finally, temperature and pressure of gases leaving the exhaust process (5'-6') are given by:

$$T_e = T_{5'} \quad (25)$$

$$p_6 = p_{5'} = p_e \quad (26)$$

The heat released by friction is computed by employing the relation [38]

$$\dot{Q}_f = \dot{m}_t \left(183 + 2.3\left(\left(\frac{N}{60}\right) - N_o\right)V_D\right) \quad (27)$$

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

Where $N_o = 30\sqrt{3/(1000V_d)}$

The heat accompanied by exhaust gases can be calculated after using:

$$\dot{Q}_{out} = \dot{m}_t[(h_4 - r.T_4) - (h_5 - r.T_5)] \quad (28)$$

The net power produced by HCCI engine:

$$\dot{W}_{net} = \dot{Q}_{in} - \dot{Q}_{out} - \dot{Q}_f \quad (29)$$

4. Energy and exergy analyses

Based on the specifications of the proposed system described above, theory leading to simultaneous applications of first and second laws of thermodynamics is employed to carry out the analysis which reveals the insights in performance of the proposed combined cycle that simultaneously generate power and cooling energy for refrigeration and air conditioning [8, 31]. The law of conservation of mass and the principle of energy balance applied to thermodynamic system operating under steady state conditions entails the following equations.

$$\Sigma\dot{m}_i - \Sigma\dot{m}_o = 0 \quad (30)$$

$$\Sigma\dot{Q} - \Sigma\dot{W} + \Sigma\dot{m}_i h_i - \Sigma\dot{m}_o h_o = 0 \quad (31)$$

Application of thermodynamic theory related to analysis of natural gas fueled HCCI engine [31] and the two-phase ejector cycle described by Wang and Yu [23], the equations leading to energy balance in the components of the natural gas fueled HCCI engine based refrigeration and air-conditioning system were derived and are reported in Table 1.

Inefficiencies in the system can be evaluated through the incorporation of exergy into the analysis. Further, inclusion of exergy leads to the determination of quality and quantity of energy

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

transformation during the change of thermodynamic state. The general expression for the exergy of flowing stream operates under the condition of steady-state and ignoring the changes in kinetic and potential energies can be presented as

$$\dot{E}_{x,D} = \sum\left(1 - \frac{T_0}{T_j}\right)\dot{Q}_{CV} - \dot{W}_{CV} + \sum \dot{m}_i e_{x,i} - \sum \dot{m}_e e_{x,e} \quad (32)$$

$\dot{E}_{x,D}$ is the exergy destruction rate and e_x is the flow exergy of the stream per unit mass and consist of: physical, chemical, potential and kinetic and is given by the following expression

$$ex = ex^k + ex^p + ex^{ph} + ex^{ch} \quad (33)$$

The physical exergy is given by

$$ex^{ph} = (h - h_o) - T_o(s - s_o) \quad (34)$$

Chemical exergy appears due to the deviations in chemical composition of the substance from the species present in the environment [31]:

$$ex^{ch} = \sum x_i(\mu_i^* - \mu_{i,o}) = -\Delta g_i^o + \sum \bar{R}T_o x_i ln\left(\frac{p}{p_{ref}}\right) \quad (35)$$

The equation which can estimate the destruction of exergy in the components of the proposed combined cycle are formulated in Table 2.

5. Criteria for the overall performance evaluation of the cycle

The thermal efficiency of the proposed system may be determined by applying first law and it can be presented as:

$$\eta_{th,comb} = \frac{\dot{W}_{net+}\,\dot{Q}_{AC} + \dot{Q}_{ref}}{\dot{m}_f \mathrm{QLHV} \eta_{cc}} \quad (36)$$

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

where, $\dot{W}_{net}$ is the net power produced by HCCI engine, $\dot{m}_f$ is fuel flow rate, Q_{LHV} is the fuel heating value of fuel (natural gas), and η_{cc} is the combustion efficiency. $\dot{Q}_{ref}$ and $\dot{Q}_{AC}$ are the cooling capacities of the refrigeration evaporator and air-conditioning evaporator, respectively, and the expression for their evaluation are presented in Table 2.

Efficiency based on exergy can be presented as

$$\eta_{ex,com} = \frac{\dot{W}_{net}+\dot{E}_{x,AC}+\dot{E}_{x,ref}}{\dot{m}_f\, ex^{ch}\, \eta_{cc}} \quad (37)$$

where, $\dot{E}_{x,AC}$ and $\dot{E}_{x,ref}$ are exergy accompanied $\dot{Q}_{AC}$ and $\dot{Q}_{ref}$ respectively, and are given by

$$\dot{E}_{x,AC} = \dot{Q}_{AC}\left(\frac{T_0-T_{AC}}{T_{AC}}\right) \quad (38)$$

$$\dot{E}_{x,ref} = \dot{Q}_{ref}\left(\frac{T_0-T_{ref}}{T_{ref}}\right) \quad (39)$$

Where, T_{ref}, T_{AC}, T_0 are the refrigeration evaporator temperature, air-conditioning evaporator temperature, and ambient temperature, respectively, and ex^{ch} is the chemical exergy of fuel.

Engineering Equation Solver (EES) and REFPROP V 9.1 was employed to evaluate the combined cycle performance [39, 40].

Combined cycle consisting of an HCCI engine and the refrigeration, and air -conditioning evaporators, following assumptions were taken to develop the thermodynamic model [22, 23, 31]:

- In the duct and heat exchanger, pressure losses are neglected.
- Engine speed was assumed to vary from 1400-2200 rpm.
- The residual gas fraction was considered to be lower because of the elevated compression ratio (f = 0.03).
- All processes are assumed to be of steady state conditions.

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

- No consideration of pressure drop in the pipes.
- For the components of two-phase ejector cooling cycle, only physical exergy of the working fluid is considered in the analysis.

The chemical exergy of fuel can be expressed as

$$e_{fuel}^{ch} = -\Delta g + \bar{R}\mathrm{T}_0\left\{x\mathrm{O}_2 ln\frac{P_{\mathrm{O}_2}^{00}}{P^0} + y\mathrm{H}_2 O ln\frac{P_{H_2O}^{00}}{P^0} - \sum_K \mathrm{x}_k ln\frac{P_k^{00}}{P^0}\right\} \quad (40)$$

The symbols appeared in above equation are defined in detail in the Ref. [27].

Description of the equations employed to model the components of the two-phase ejector which is considered in the proposed configuration is formulated in Ref. [23].

6. Results and discussion

The natural gas fueled HCCI engine integrated with the two-phase ejector cooling cycle is investigated and explored parametrically. R717, R290, and R600a were employed as refrigerant in the two-phase ejector. Properties of these refrigerants are presented in Table 3. The effects of changing the equivalence ratio, engine speed, refrigerant condenser temperature, efficiency of ejector nozzle, refrigeration evaporator temperature, and air conditioning evaporator temperature is investigated on the performance of proposed system. Properties utilized in the study of HCCI engine based combined cycle for its base line operation are shown in Table 4 and they are in line with the reported values of the literature.

In Fig. 3, impact of varying the equivalence ratio on the thermal efficiency of combined cycle is observed for the use of three different refrigerants in the two-phase ejector. The energy supplied to the engine is determined by the equivalence ratio. Increase in equivalence ratio results in the greater release of energy which elevates the in-cylinder pressure and temperature and this in turn

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

increase the power generated by the engine from 129 kW to 135.8 kW when an equivalence ratio is promoted from 0.3 to 0.9. Further, any increment in equivalence ratio leads to the elevation of temperature of catalytic convertor exhaust which expanded in the turbine to generate power in order to drive the compressors employed in the combined cycle. The power supplied to drive the compressor of bottoming cooling cycle (W_{C2}) is raised from 15.69 kW to 16.57 kW when the equivalence ratio is increased from 0.3 to 0.9. This in turn increases the refrigeration capacity Q_{ref} from 37.43 kW to 39.52 kW and the cooling capacity of air conditioner Q_{AC} from 9.91 kW to 10.47 kW. Enhancement in the performance of HCCI engine along with the simultaneous increase in the output of two-phase ejector cooling cycle results in the significant increase of the efficiency of combined cycle. Variation of equivalence ratio from 0.3 to 0.9 causes an increase of thermal efficiency of combined cycle from 60.05% to 63.26% when R600a is the refrigerant and from 57.87% to 60.89% when R717 is employed as the refrigerant, and for the application of R290 as the refrigerant the efficiency rises from 57.67% to 60.73%. This deviation is observed due to the fact that the mass flow rate of refrigerant enters the evaporator is found highest in the R600a and lowest in R717 and hence the rate of refrigeration produced by R600a operated bottoming cycle is significantly higher due to which the thermal of the cycle is also higher. In case of ammonia (R717), the mass flow rate of refrigerant is less but the enthalpy of vaporization is found to be greater. On the other hand, in case of propane (R290), the mass flow rate is little higher but enthalpy of vaporization is less. Since the enthalpy of vaporization dominates over the mass flow rate, therefore, the thermal efficiency of R717 operated combined cycle is slightly higher than R290 operated cycle.

Since exergy efficiency pertaining to HCCI engine is defined as the ratio of power exergy to exergy of fuel supplied, therefore, exergy efficiency is lower than the corresponding thermal

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

efficiency of HCCI engine because fuel exergy is found larger than the fuel energy. Change in equivalence ratio from 0.3 to 0.9 entails an increase in the exergy efficiency of HCCI engine from 40.15% to 42.12% while its thermal efficiency increased from 47.44% to 49.94% as shown in Fig.4. It is interesting to see the exergy efficiency of combined cycle (power and cooling) is lower than the exergy efficiency of HCCI engine. This is because in the proposed configuration the net power output of HCCI engine alone is the HCCI power plus the net power output of the turbocharger (W_T-W_{C1}). In the absence of the bottoming cooling cycle, the net power of turbocharger will be added to engine power and hence the net power output of HCCI will increase. In case of the employment of bottoming cooling cycle, the net power of turbocharger will be utilized to drive the cooling cycle whose evaporators are producing the exergy of refrigeration and exergy of air conditioning but these exergies are considerably less than turbocharger net output.

Figs. 5 and 6 shows the impact of change in engine speed on the performance of HCCI engine and combined cycle, and both are found to be largely dependent on engine speed. This is because at higher engine speed fuel-air mixing is improved and the charge-flow intensity is increased. Moreover, the increase in engine speed decreases the loss of heat to cylinder wall and loss in exhaust gases. Due to these factors, the thermal efficiency of HCCI engine increase as engine speed increases (Fig. 5). Since the exergy accompanied by the loss of heat to cylinder wall and the engine exhaust gases are also reduced with the engine speed, therefore, its exergy efficiency also increases. The thermal efficiency of combined cycle is higher than the exergy efficiency of HCCI engine because of the reasons explained for the trends obtained in Fig. 3. In addition, Fig. 5 shows that thermal efficiency of combined cycle increases because heat supplied to the combined cycle decreases due to reduced heat losses at higher engine speed. Since the exergy associated with heat

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

input also decrease as the engine speed increases which results in increase in the exergy efficiency of the combined cycle as depicted in Fig. 6.

Fig. 7 shows the effect of change in the AC evaporator temperature (T_{AC}) on the thermal and exergy efficiencies. The temperature of AC evaporator is irrelevant to net power output of HCCI engine but it has an impact on the cooling capacity of AC evaporator (Q_{AC}) and also on the cooling capacity of the refrigeration evaporator (Q_{ref}) of the bottoming cycle. As T_{AC} increases, the thermal efficiency of the combined cycle goes up in case of the use of R600a and R717 and in case of R290 applications as working fluid, and it shows a downward trend. This is because in case of all three refrigerants, Q_{AC} increases with the increase of T_{AC}. Further, increase in T_{AC} raises the enthalpy of refrigerant enter the ejector which result in the increase of its velocity at the nozzle exit, causing to extract more fluid from the refrigerant evaporator which results in the rise of Q_{ref} in case of R717 and R600a fluids while in case of R290 applications, a little decrement in Q_{ref} is observed. Since the contribution of Q_{AC} and Q_{ref} towards the total energetic performance of the combined cycle is quite small compared to net power output of the HCCI engine, therefore, a considerable increase in T_{AC} is resulting in the insignificant variation of the thermal efficiency of combined cycle. The exergy efficiency of the combined cycle is lower than its thermal efficiency, because exergy associated with Q_{AC} and Q_{ref} is significantly less than the magnitudes of Q_{AC} and $Q_{ref.}$ The variation is quite insignificant because in case of R717 and R600a fluids Q_{AC} decreases while Q_{ref} increases and in case of R290 applications both Q_{AC} and Q_{ref} are found to be decreasing with the increase of T_{AC}.

Fig. 8 illustrates the effect of change of refrigeration evaporator temperature (T_{ref}) on the thermal and exergy efficiencies of the combined cycle. The T_{ref} is unrelated to HCCI engine power output. Thus, the engine power output and the power exergy are remained unchanged. An increase in T_{ref}

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

increases the refrigeration capacity Q_{ref} due to which the energetic outcome of the combined cycle increased and hence its thermal efficiency increases. The decrement in the exergy efficiency of the cycle is so small because the rate of decrease in the exergy of refrigeration reduces with the rise of refrigeration temperature. It is seen as T_{ref} increases the thermal efficiency increases from 60.04% to 63.52% for R600a operated cycle and it is increased from 57.58% to 61.07% for R717 operated cycle when T_{ref} rises from -25^0C to -15^0C.

As depicted in Fig. 9, thermal efficiency of combined cycle declines markedly with growing $T_{cond.}$ The condenser temperature is irrelevant to net power output of the engine, so the turbine and compressors power remain unchanged. For greater values of T_{cond}, large enthalpy difference across the compressor is determined. For the fixed compressor power, lower mass flow rate of refrigerant enters the AC evaporator at state (13) which reduces the cooling capacity. Reduced mass flow rate of primary stream at state (14) causes the corresponding lowering of secondary fluid flow rate and consequent drop in cooling capacity of refrigeration evaporator. Simultaneous reduction in the cooling capacity of AC evaporator and refrigeration evaporator causes combined cycle efficiency to decrease considerably with the rise of condenser temperature. Since combined cycle thermal efficiency is weakly prevailed by Q_{AC} and $Q_{ref,}$ therefore, thermal efficiency exhibits a small reduction with increasing T_{cond} for all of the refrigerants. It is further seen in Fig. 9 that exergy efficiency of the combined cycle shares a common characteristic: decreasing as T_{cond} increases. It is determined that as T_{cond} rises, a very slight decline in the exergy efficiency for all of the refrigerants is observed. This is because the exergy accompanied by cooling water increases with the increase of condenser temperature which in turn reduces the exergy accompanied by Q_{AC} and Q_{ref}. Computation of results reveal that $Q_{ref.}$ decreases from 44 kW to 33 kW and the Q_{AC} decreases from 11 kw to 9 kW while the exergy accompanied by Q_{ref} decreases 8.1 to 6.2 kw and the exergy

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

associated with Q_{AC} decreases from 0.28 kw to 0.21 kW when T_{cond} increases from 30^0C to 40^0C. Since the exergy associated with Q_{AC} and Q_{ref} are quite small compared to Q_{AC} and $Q_{ref,}$ therefore, the exergy efficiency despite of showing a similar trend appears to change insignificantly contrary to a significant change in thermal efficiency of the cycle. For R717 fluid, the thermal efficiency of the cycle decreases from 61.69% to 57.41% and exergy efficiency also decreases marginally from 39.01% to 38.57% consequent to the rise T_{cond} from 30^0C to 40^0C.

The influence of the nozzle efficiency on performance of the combined cycle is shown in Fig. 10. Both thermal and exergy efficiencies of the combined cycle are found to be increasing with the rise of nozzle efficiency. Enhancement in nozzle efficiency of the ejector, enhances the velocity of the motive flow stream at the exit of nozzle which results in the simultaneous increase of the mass flow rate of primary fluid and the entrained secondary fluid from the refrigeration evaporator. This leads to the increase of cooling capacities of the AC evaporator and refrigeration evaporator of the system. Due to this reason, increase in nozzle efficiency, thermal efficiency of the combined cycle increases slightly. As the exergy associated by the cooling capacities of both the evaporator also increases, therefore, the exergy efficiency of the combined cycle also shares the same trend of increasing with the increase of nozzle efficiency. The same trend is observed for all the three refrigerants. An increasing trend is found for the application of all three refrigerants but the change in refrigerant does not exhibit a significant variation in the thermal and exergy efficiencies of the combined cycle.

After carrying out the energetic performance study, the distribution of energy possessed by natural gas which is supplied to the cycle is computed in terms of energy output and energy loss at the base case and is illustrated in Fig. 11. The results show that for R600a refrigerant, from 305 kW of energy input the energetic output of the cycle appears as; net power produced by HCCI engine

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

is 132.4 kW, cooling capacity of refrigeration evaporator is 42.75 kW, cooling capacity of AC evaporator is 12.93 kW. Engine heat loss is 108.94 kW, heat loss via condenser is 66.78 kW, and heat loss via thermal exhaust to ambient is 52.56 kW. Employment of a different refrigerant in the bottoming cycle does not alter the engine power output and energy losses but it has an impact on the cooling capacities and condenser heat loss. It is found that in case of R717 refrigerant, Q_{ref} becomes 38.48 kW, Q_{AC} becomes 10.19 kW, and condenser heat loss becomes 59.77 kW. For R290 refrigerant, Q_{ref} appears as 37.22 kW, Q_{AC} is 12.93 kW, and condenser heat loss becomes 59.30 kW.

Fig. 12 presents the breaking down of exergy of natural gas in the combined cycle at the base case. The total input exergy to the combined cycle is 361 kW (100%). It is evident that out of 100% exergy input, in case of R717 operated combined cycle, 139.79 kW (38.72%) is the total exergy output. Exergy destruction and losses are found as 164.21 kW (45.49%) and 57 kW (15.79%), respectively. For R290 refrigerant, exergy output is 139.58 kW (38.67%), exergy destruction is 163.98 kW (45.42%), and exergy loss is 57.44 kW (15.91%). For R600a, exergy produced is 140.65 kW (38.96%), exergy destruction is 163.75 kW (45.36%), and exergy loss is 56.6 kW (15.67%). This data clearly show that change in refrigerant from R717 to R600a, and to R290, minorly influence the percentages of exergy produced, exergy destroyed, and exergy loss.

A detailed distribution of exergy destruction and exergy loss in each component of the combined cycle is illustrated in Fig. 13 as an appropriate exergy analysis provide a better understanding of the irriversibility of the proposed combined cycle. Detailed analysis reveals HCCI engine as the major exergy destructive component where 113.3 kW (31.38%) of the exergy is destroyed. Moreover, 36.84 kW (10.21%) of exergy loss in detected in HCCI engine that is the exergy of heat losses from cylinder walls to the ambient. Exergy destruction and exergy losses in HCCI

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

engine are significantly high because of the irrevesible nature of the combustion process. Next highest destruction of exergy of 31.47 kW (8.71%) is found in the catalytic convertor. This is because during this process, the combustion gases react with the material of convertor that leads to a chemical reaction which is the main source of entropy generation. Exergy destruction in the other components of the system does not exhibit significantly because these components does not directly utilize fuel exergy but instead heat exchanging process between the two fluid streams are involved that contribute to entropy generation. It is indicating that that there is no potential for improvement under the operating parameters considered. Further, loss of exergy from the fluid stream of the combined cycle to the ambient is also computed which found to be varies with the change of refrigerant in the two-phase ejector cooling cycle. It is shown that for the application of R600a, R717, and R290 the losses of exergy are found as 19.76 kW (5.47%), 20.16 kW (5.58%), and 20.6 kW (5.70%), respectively. This shows that exergy flow loss does not change appreciably with the change of refrigerant assumed in the analysis of the proposed combined cycle.

7. Conclusion

In the current study, a new cogeneration cycle for the simultaneous production of power and cooling energy is proposed and investigated theoretically from energy and exergy point of views. The proposed combined cycle is the combination of a natural gas fueled HCCI engine and an exhaust gas operated turbine driven two-phase ejector cycle which simultaneously refrigerate the thermal loads of low temperature refrigeration and air conditioning. A thermodynamic analysis is carried out and the computed results are commented upon to reveal a deeper insight to the effect of combined cycle operational parameters (equivalence ratio, engine speed, condenser temperature, evaporators temperature, and ejector nozzle efficiency) on selected performance

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

parameters of the cycle. Thermal and exergy efficiencies of the combined cycle are evaluated and the comparative results are presented.

Following can be considered as the main concluding remarks which inferred from this study:

- As the equivalence ratio increases (0.3-0.9), the thermal efficiency of the HCCI engine is increased from 47.44% to 49.94%. Thermal efficiency of the combined cycle display a similar trend of increasing and rises from 60.05% to 63.26% in case of R600a operated cooling cycle and this increasing trend is observed for all the considered refrigerants.
- Increase in equivalence ratio from 0.3 to 0.9 causes an increase in HCCI engine exergy efficiency from 42.15% to 42.12% and the combined cycle exergy efficiency for R600a application is found to be increased from 38.02% to 39.89%, respectively. Variation of exergy efficiency display a similar charactreistic for other two refrigerants R290 and R717.
- Increasing of engine speed in the range of 1400-2200 r.p.m. leads to an increasing trend in both the efficiencies of HCCI engine and the combined cycle. For R600a refrigerant, $\eta_{th,cc}$ increases from 58.02% to 65.58% and the $\eta_{ex,cc}$ increases from 37% to 41%, respectively.
- The influence of increase in condenser temperarture (T_{cond}) leads to a decreasing trend for the thermal efficiency of combined cycle. It is found that for R717 $\eta_{th,cc}$ decreases from 61.69% to 57.41%, for R290 decreased from 60.99% to 57.69%, and for R600a it is decreased from 63.85% to 59.87%, respectively, when T_{cond} rises from 30^0C to 40^0C.
- Energetic analysis of the cycle carried out show that employment of a different refrigerant in the bottoming cycle does not alter the engine power output and energy losses but it has an impact on the cooling capacities and condenser heat loss.

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

- Distribution of fuel exergy results show that out of 100% exergy input, in case of R717 operated combined cycle, 139.79 kw (38.72%) is the total exergy output and 164.21 kW (45.49%) and 57 kW (15.79%), are the values for exergy destruction and exergy losses. Further, it is determined that change in refrigerant from R717 to R600a, and to R290, minorly influence the percentages of exergy produced, destroyed, and loss.
- Among all the components of the combined cycle HCCI engine had the main contribution in exergy destruction (31.38%) and exegy loss (10.21%) for all the refrigerants considered.

Nomenclature

A_{ch}	surface of cylinder head (m^2)
A_{cyl}	surface of the cylinder (m^2)
C_p	specific heat at constant pressure (kJ/kg)
$\dot{E}_x$	exergy rate (kW)
e_x^{ch}	chemical exergy (kJ/kg)
e_x^{ph}	physical exergy (kJ/kg)
e_x	specific flow exergy (kJ/kg)
f	fraction of residual gases
g	Gibbs function (kJ/kg)
$\bar{h}_f^\circ$	enthalpy of formation on molar basis (kJ/kmol)
H	enthalpy (kW)
h	specific enthalpy (kJ/kg)

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

$\dot{m}$	mass flow rate (kg/s)
M	molar mass (kg/kmol)
n	number of moles, polytropic index
p_{mot}	pressure at motored condition (kPa)
p	pressure (kPa)
$\dot{Q}$	rate of heat transfer (kW)
R	molar gas constant (J/K mol)
s	specific entropy [kJ/kg-K]
T	temperature [K]
V	volume (m^3)
$\dot{W}$	work output [kW]
x	mass fraction
y	mole fraction

Greek symbols

η	efficiency [%]
$\emptyset$	equivalence ratio
μ	chemical potential
ω	mean gas velocity

Subscript

a	air

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

avr	average
AC	air-conditioner
BT	brake thermal
C1	compressor 1
C2	compressor 2
CC	combined cycle, catalytic convertor
cat	catalytic
ch	chemical
comb	combustion
comp	compression
Cond	condenser
D	destruction
Eje	ejector
EV1	expansion valve 1
EV2	expansion valve 2
ex	exergy
exp	expansion
f	fuel
FM	fuel mixer
HX	heat exchanger

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

i	inlet
LHV	lower heating value
o	outlet
ph	physical
ref	refrigerator, reference
Reg	regenerator
Sep	separator
surr	surrounding
T	turbine
th	thermal

Abbreviations

HCCI	Homogeneous charge compression ignition
ARS	Absorption refrigeration system
ERC	Ejector refrigeration cycle

References

[1] Armstead, J.R., and Miers, S.A., 2014, "Review of waste heat recovery mechanisms for internal combustion engine," ASME J. Thermal Sci. Eng. Appl., 6 (1), pp. 014001. https://doi.org/10.1115/1.4024882

[2] Srinivas, K.K., Mago, P.J., and Krishnan, S.R., 2010, "Analysis of exhaust waste heat recovery from a dual fuel low temperature combustion engine using an organic Rankine cycle," Energy, 35(6), pp. 2387-2399, https://doi.org/10.1016/j.energy.2010.02.018

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

[3] Manzela, A.A., Hanriot, S.M., Cabezas-Gomez, L., and Sodre, J.R., 2010, "Using engine exhaust gas as energy source for an absorption refrigeration system," Appl. Energy, 87, pp. 1141-1148.

[4] Qin, F., Chen, J., Lu, M., Chen, Z., Zhou, Y., and Yang, K., 2007, "Development of a metal hydride refrigeration system as an exhaust gas driven automobile air conditioner," Renewable Energy, 32, pp. 2034-2052.

[5] Lang, L., Colonna, P., and Almbauer, R., 2013, "Assessment of waste recovery from a heavy-duty truck engine by means of an ORC turbogenerator," ASME J. Eng. Gas Turbines Power, 135 (4), pp. 042313. https://doi.org/10.1115/1.4023123

[6] Dolz, V., Novella, R., Garcia, A., and Sanchez, J., 2012, "HD Diesel engine equipped with a bottoming Rankine cycle as a waste heat recovery system. Part 1: study and analysis of the waste energy," Appl. Therm. Eng., 36, pp. 269-278. https://doi.org/10.1016/j.applthermaleng.2011.10.025

[7] Yu, Z., Han, J., Liu, H., and Zhao, H., 2014, "Theoretical Study on a Novel Ammonia-Water Cogeneration System with Adjustable Cooling to Power Ratios," Appl Energy, 122, pp. 53-61.

[8] Chen, Y., Han, W., Jin, H., 2017, "Investigation of an ammonia-water combined power and cooling system driven by the jacket water and exhaust gas heat of an internal combustion engine," Int. Journal of Refrigeration, 82, pp. 174-188.

[9] Yu, W., Xu, Y., Wang, H., Ge, Z., Wang, J., Zhu, D., and Xia, Y., 2020, "Thermodynamic and thermoeconomic performance analyses and optimization of a novel power and cooling cogeneration system fueled by low-grade waste heat," Appl. Therm. Eng., 179, pp. 115667. https://doi.org/10.1016/j.applthermaleng.2020.115667

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

[10]Ruangtrakoon, K., and Aphornratana, S., 2019, "Design of steam ejector in a refrigeration application based on thermodynamic performance analysis," Sustainable Energy Technologies and Assessments, 31, pp. 369-382. https://doi.org/10.1016/j.seta.2018.12.014

[11]Abdulateef, J.M., Sopian, K., Alghoul, M.A., and Sulaiman, M.Y., 2009, "Review on solar-driven ejector refrigeration technologies," Renew. Sustain. Energy Rev., 13, pp. 1338-1349.

[12]Alexis, G.K., and Karayiannis, E.K., 2005, "A solar ejector cooling system using refrigerant R134a in the Athens area," Renew. Energy, 30 (9), pp. 1457–1469. https://doi.org/10.1016/j.renene.2004.11.004

[13]Bourhan, T., Aiman, A., Saja, A., 2015, "Performance study of ejector cooling cycle at critical mode under superheated primary flow," Energy Convers. Manage., 94, pp. 300-310.

[14]Chandra, V.V., and Ahmed, M.R., 2014, "Experimental and computational studies on a steam jet refrigeration system with constant area and variable area ejectors," Energy Convers. Manage., 79, pp. 377-386.

[15]Yılmaz, A., and Aktas, A.E., 2019, "Comparative Analysis of Ejector Refrigeration System Powered with Engine Exhaust Heat using R134a and R245fa," European Mechanical Science, 3(1), pp. 13-17. https://doi.org/10.26701/ems.430831

[16]Jaruwongwittaya, T., and Chen, G., 2012, "Application of two stage ejector cooling system in a bus," Energy Procedia, 14 pp. 187-197. https://doi.org/10.1016/j.egypro.2011.12.916

[17]Xia, J., Wang, J., Lou, J., Zhao, P., and Dai, Y., 2016, "Thermo-economic analysis and optimization of a combined cooling and power (CCP) for engine waste heat recovery,"

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

Energy Convers. Manage., 128, pp. 303-316. https://doi.org/10.1016/j.enconman.2016.09.086

[18]Kornhauser, A.A., 1990, "The use of an ejector as a refrigerant expander," International Refrigeration and Air Conditioning Conference. Paper 82. http://docs.lib.purdue.edu/iracc/82

[19]Chaiwongsa, P., and Wongwises, S., 2007, "Effect of throat diameters of the ejector on the performance of the refrigeration cycle using a two-phase ejector as an expansion device," Int. J. Refrigeration, 30, pp. 601-608.

[20]Pottker, G., Guo, B., and Hrnjak, P.S., 2010, "Experimental investigation of an R410A vapor compression system working with an ejector," International Refrigeration and Air Conditioning Conference, Paper 1135. http://docs.lib.purdue.edu/iracc/1135

[21]Lawrence, N.D., 2012, "Analytical and experimental investigation of two-phase ejector cycles using low-pressure refrigerants (MS Thesis)," University of Illinois, Urbana, IL, USA. http://hdl.handle.net/2142/42337

[22]Lawrence, N., and Elbel, S., 2013, "Theoretical and practical comparison of two-phase ejector refrigeration cycles including first and second law analysis," Int. J. Refrigeration, 36, pp. 1220-1232.

[23]Wang, X., and Yu, J., 2016, "An investigation on the component efficiencies of a small two-phase ejector," Int. J. Refrigeration, 71, pp. 26-38.

[24]Zegenhagen, M.T., and Ziegler, F., 2015, "Feasibility analysis of an exhaust gas waste heat driven jet-ejector cooling system for charge air cooling of turbocharged gasoline engines," Appl. Energy, 160 pp. 221-230. https://doi.org/10.1016/j.apenergy.2015.09.057

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

[25] Unal, S., 2015, "Determination of the ejector dimensions of a bus air-conditioning system using analytical and numerical methods," Appl. Therm. Eng., 90, pp. 110-119. https://doi.org/10.1016/j.applthermaleng.2015.06.090

[26]Zhang, H., Wang, L., Jia, L., and Wang, X., 2018, "Performance investigation of automobile waste heat recovery system for ejector refrigeration cycle" 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), DOI: 10.1109/ICIEA 2018.8397750.

[27]Siddiqui, M.A., Khaliq, A., and Kumar, R., 2021, "Proposal and analysis of a novel cooling-power cogeneration system driven by the exhaust gas heat of HCCI engine fueled by wet-ethanol," Energy, 232, pp. 120954. https://doi.org/10.1016/j.energy.2021.120954

[28]Yu, W., Wang, H., Ge, Z., 2021, "Comprehensive analysis of a novel power and cooling cogeneration system based on organic Rankine cycle and ejector refrigeration cycle," Energy Convers. Manage., 232, pp. 113898. https://doi.org/10.1016/j.enconman.2021.113898

[29]Khaliq, A., Islam, S., and Dincer, I., 2019, "Energy and exergy analyses of a HCCI engine-based system running on hydrogen enriched wet-ethanol fuel," Int J Exergy, 28 (1), pp. 72-95. https://doi.org/10.1504/IJEX.2019.097272

[30]Parvez, M., and Khaliq, A., 2014, "Exergy analysis of syngas fuelled cogeneration cycle for combined production of power and refrigeration," Int J Exergy, 14(1), pp. 1-21

[31]Djermouni, M., and Ouadha, A., 2014, "Thermodynamic analysis of an HCCI engine based system running on natural gas," Energy Convers. Manage., 88, pp. 723-731.

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

[32]Khaliq, A., and Trivedi, S.K., 2012, "Second law assessment of a wet ethanol fuelled HCCI engine combined with organic Rankine cycle," ASME J Energy Resour Technol., 134, pp. 1-12.

[33]Frias, J.M., Aceves, S.M., and Flowers, D.L., 2007, "Improving ethanol life cycle energy efficiency by direct utilization of wet-ethanol in HCCI engines," ASME J. Energy Resour. Technol, 129(4), pp. 332-337. https://doi.org/10.1115/1.2794768

[34]Zheng, Z.Q., Yao, M.F., Chen, Z., and Zhang, B., 2004, "Experimental study on HCCI combustion of dimethyl ether (DME)/methanol dual fuel," SAE Paper 2004-01-2993.

[35]Mo, Y., 2002, "HCCI heat release rate and combustion efficiency: a coupled kiva multi-zone modeling study," PhD Thesis, University Michigan, USA.

[36]Soyhan, H.S., Yasar, H., Head, B., Kalghatgi, G.T., and Sorusbay, C., 2009, "Evaluation of heat transfer correlations for HCCI engine modeling," Appl Therm Eng., 29, pp. 541-549.

[37]Heywood, J.B., 1998, "Internal Combustion Engine Fundamentals," New York: McGraw Hill.

[38]Olof, E., 2002, "Thermodynamic simulation of HCCI engine systems," PhD Thesis, Lund University, Sweden.

[39]Klein, S.A., 2012, "Engineering Equation Solver (EES) for Microsoft Windows operating systems: Academic professional version: F-chart software," Madison, W.I. (Available online http://www.fchart.com)

[40]REFPROP, 2013, "NIST Reference Thermodynamic and Transport Properties," version 9.1.

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

FIGURE CAPTIONS:

Fig. 1. Natural gas operated power and cooling combined cycle

Fig. 2: PV diagram

Fig. 3. Effect of equivalence ratio on the thermal efficiency of HCCI engine and combined cycle

Fig. 4. Effect of equivalence ratio on the exergy efficiency of HCCI engine and combined cycle

Fig. 5. Effect of engine speed on the thermal efficiency of HCCI engine and the combined cycle

Fig. 6. Effect of engine speed on the exergy efficiency of HCCI engine and the combined cycle

Fig. 7. Effect of AC evaporator temperature on the thermal and exergy efficiencies of combined cycle

Fig. 8. Effect of refrigeration evaporator temperature on the thermal and exergy efficiencies of combined cycle

Fig. 9 Variation of thermal and exergy efficiencies of combined cycle with condenser temperature

Fig. 10. Variation of thermal and exergy efficiencies of combined cycle with nozzle efficiency

Fig. 11. Distribution of fuel energy in natural gas operated HCCI engine based combined cycle

Fig. 12. Distribution of fuel exergy input in the produced, exergy destroyed and exergy loss in the proposed combined cycle

Fig. 13. Breaking down of the fuel exergy input in the components of the proposed combined cycle

TABLE CAPTIONS:

Table 1: Component wise energy balance equations for natural gas operated HCCI engine based refrigeration and air-conditioning system

Table 2: Component wise exergy balance equations for natural gas operated HCCI engine based refrigeration and air-conditioning system

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

Table 3: Characteristic of Refrigerants

Table 4: Input data used for the computation of results of the natural gas operated HCCI engine based refrigeration and air-conditioning system

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

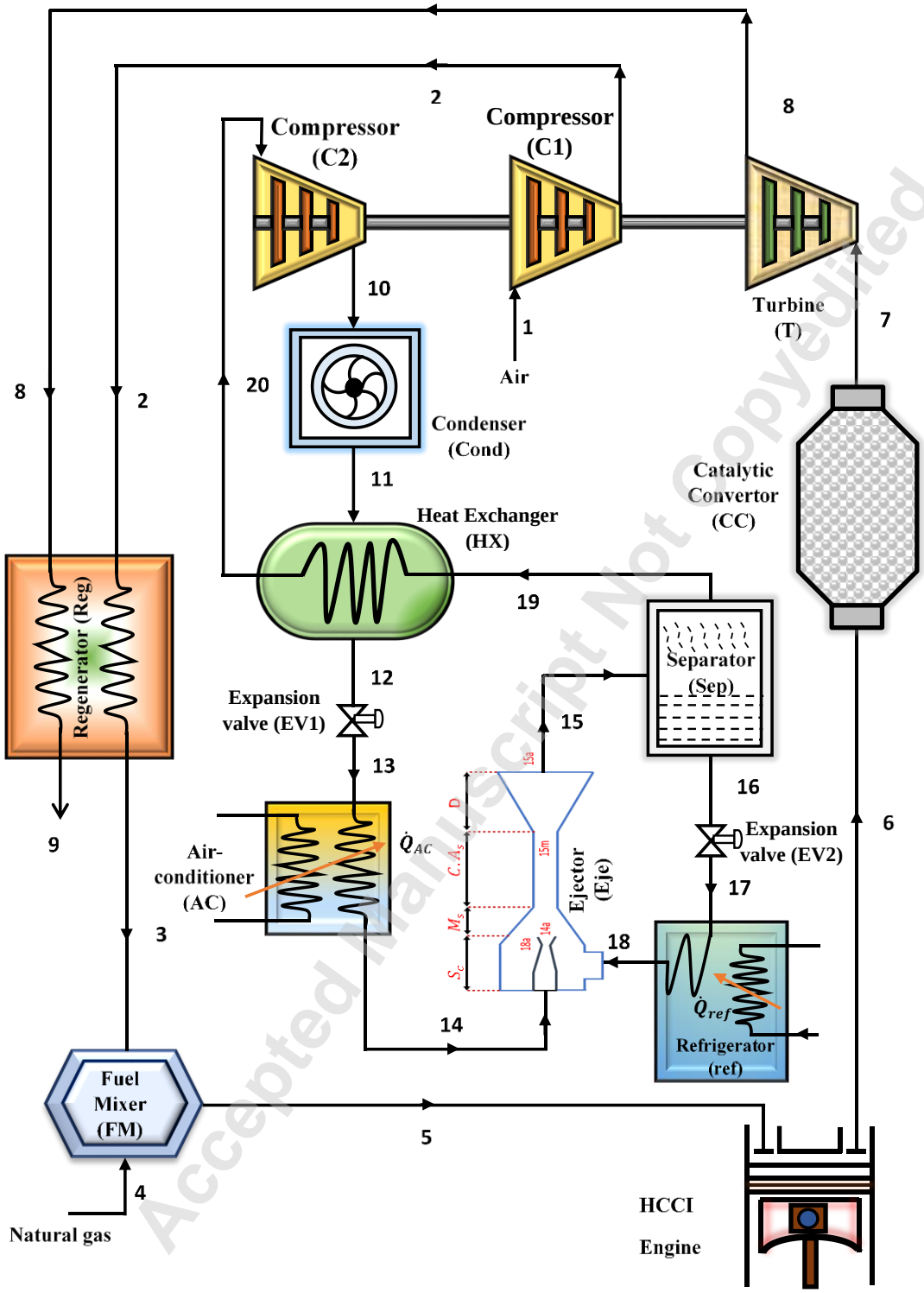


Fig. 1. Natural gas operated power and cooling combined cycle

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

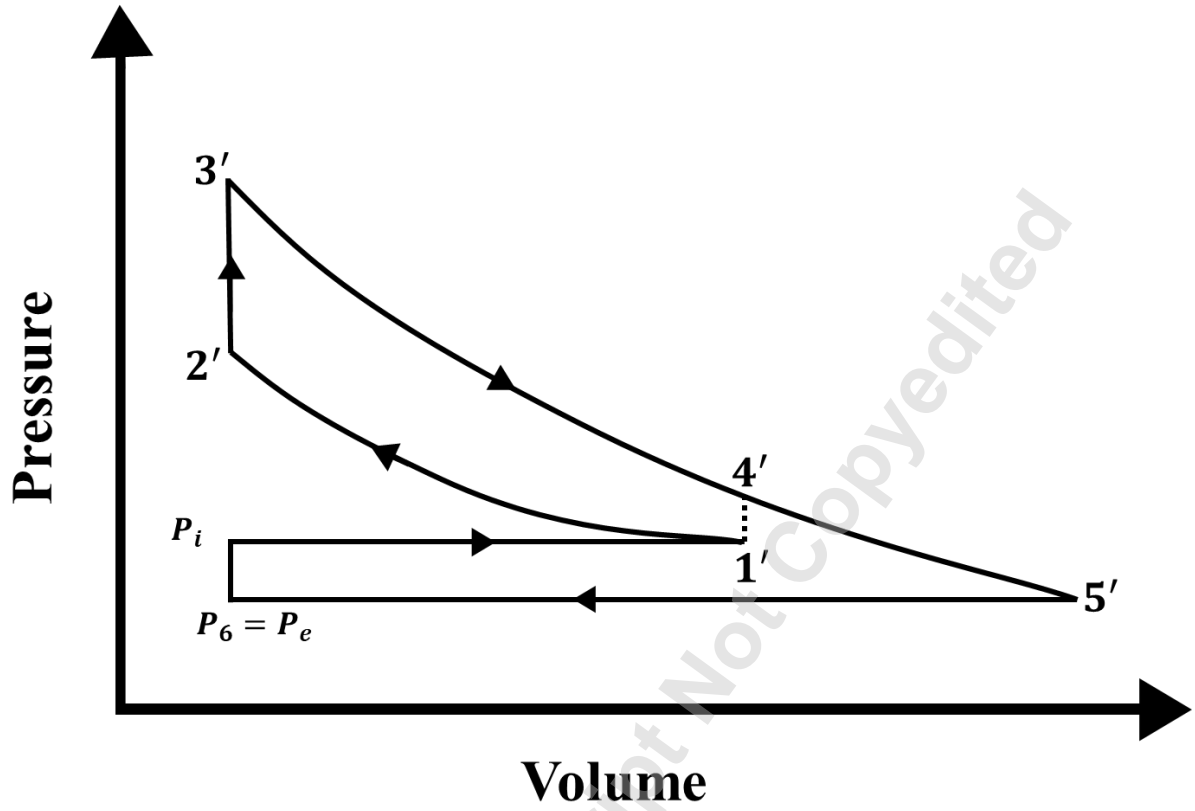


Fig. 2: PV diagram

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

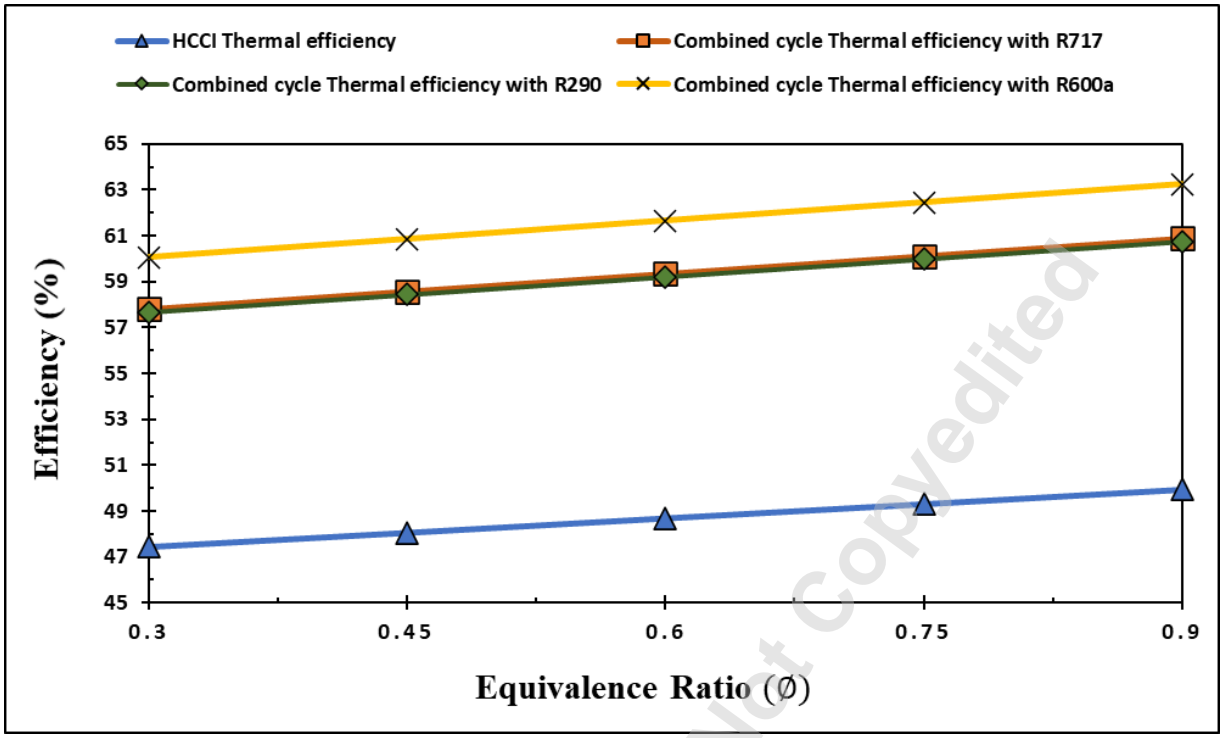


Fig. 3. Effect of equivalence ratio on the thermal efficiency of HCCI engine and combined cycle

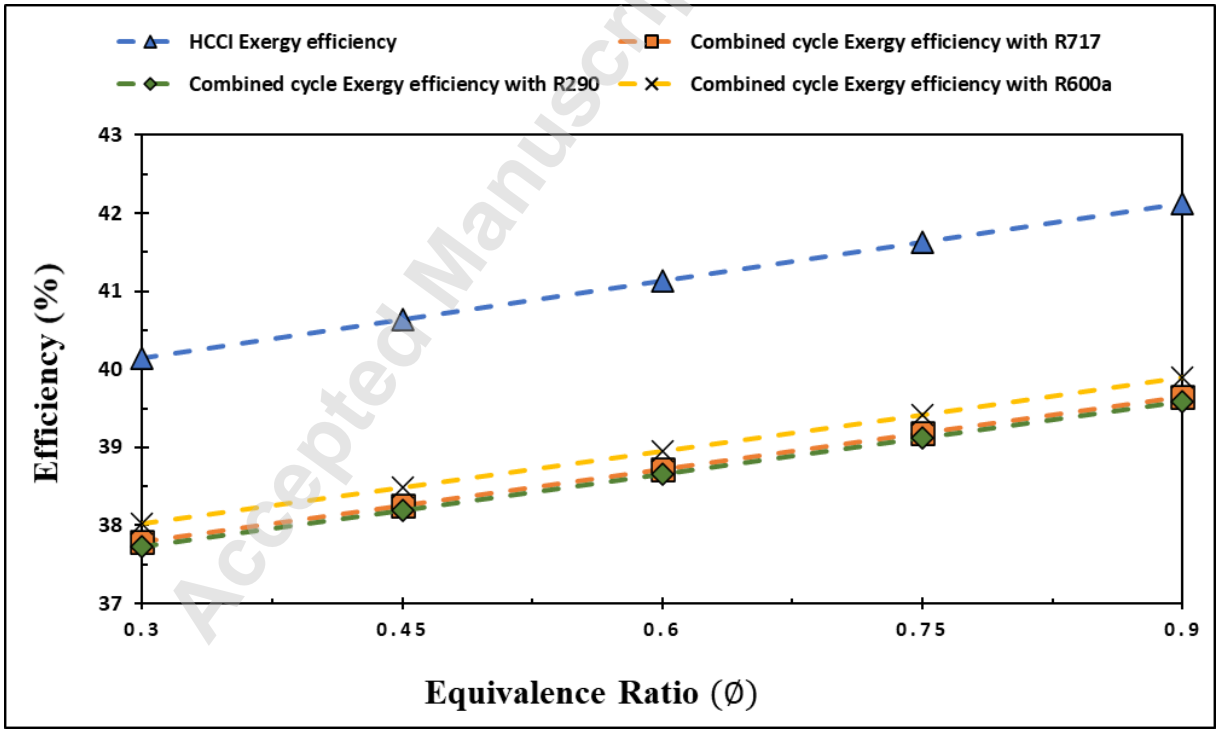


Fig. 4. Effect of equivalence ratio on the exergy efficiency of HCCI engine and combined cycle

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

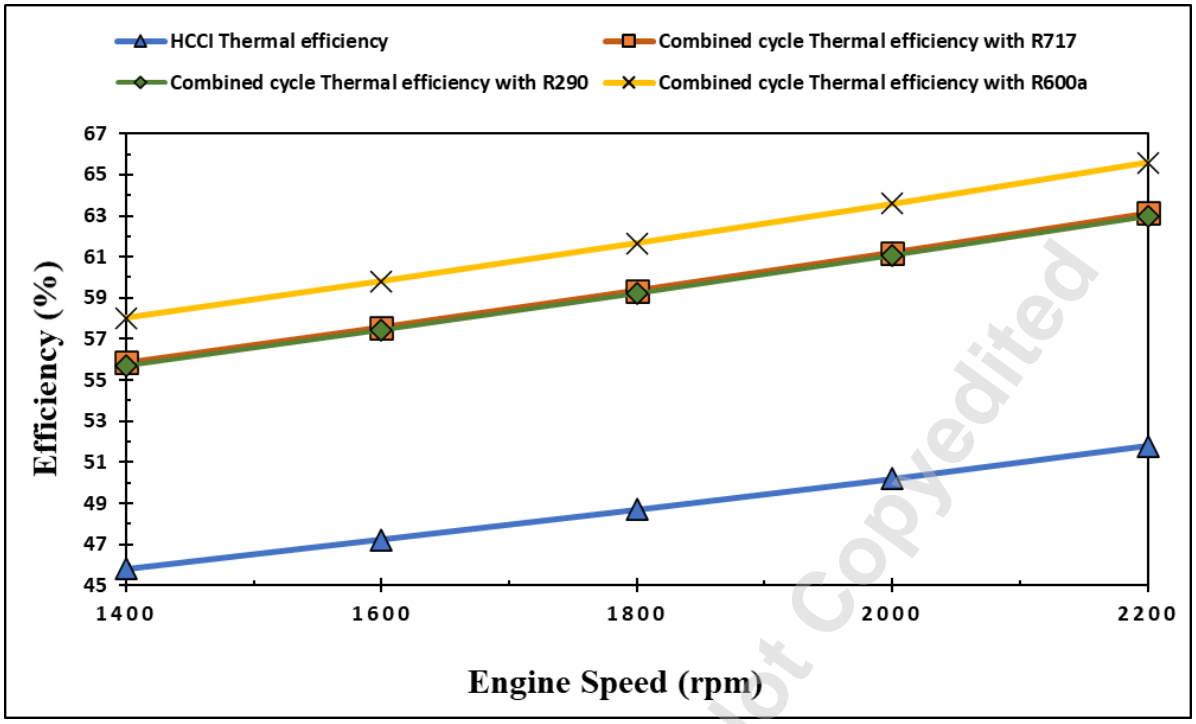


Fig. 5 Effect of engine speed on the thermal efficiency of HCCI engine and the combined cycle

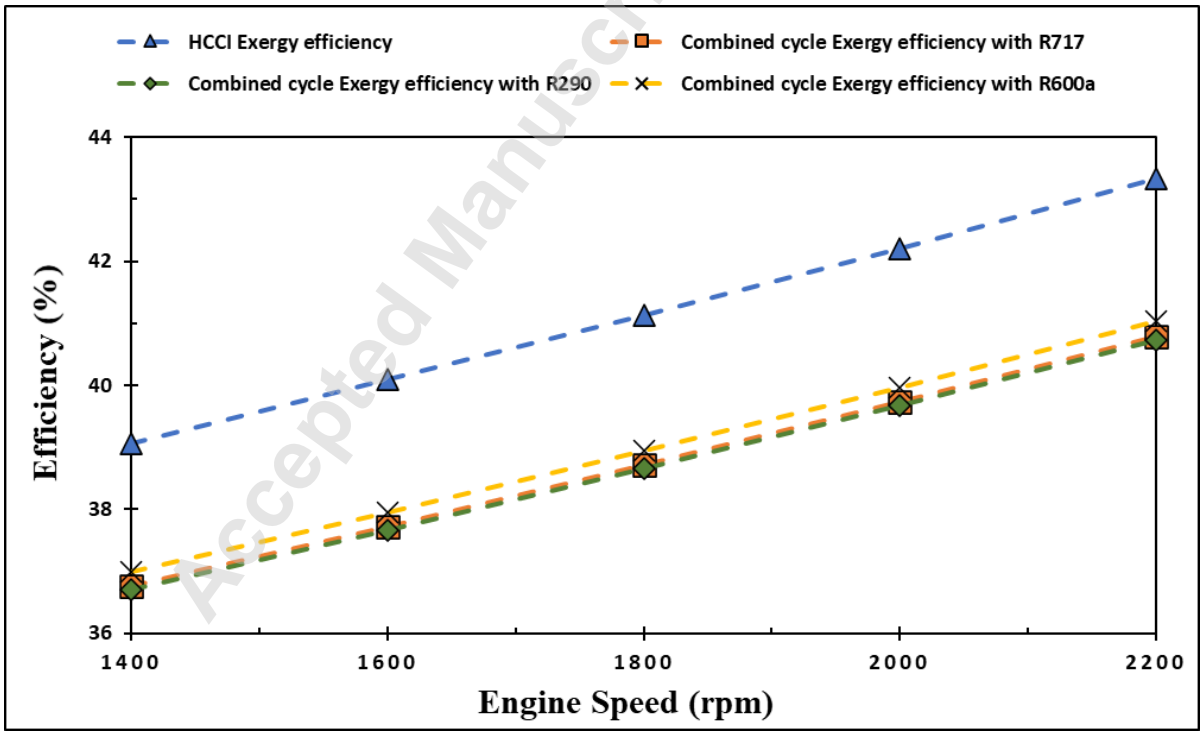


Fig. 6. Effect of engine speed on the exergy efficiency of HCCI engine and the combined cycle

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

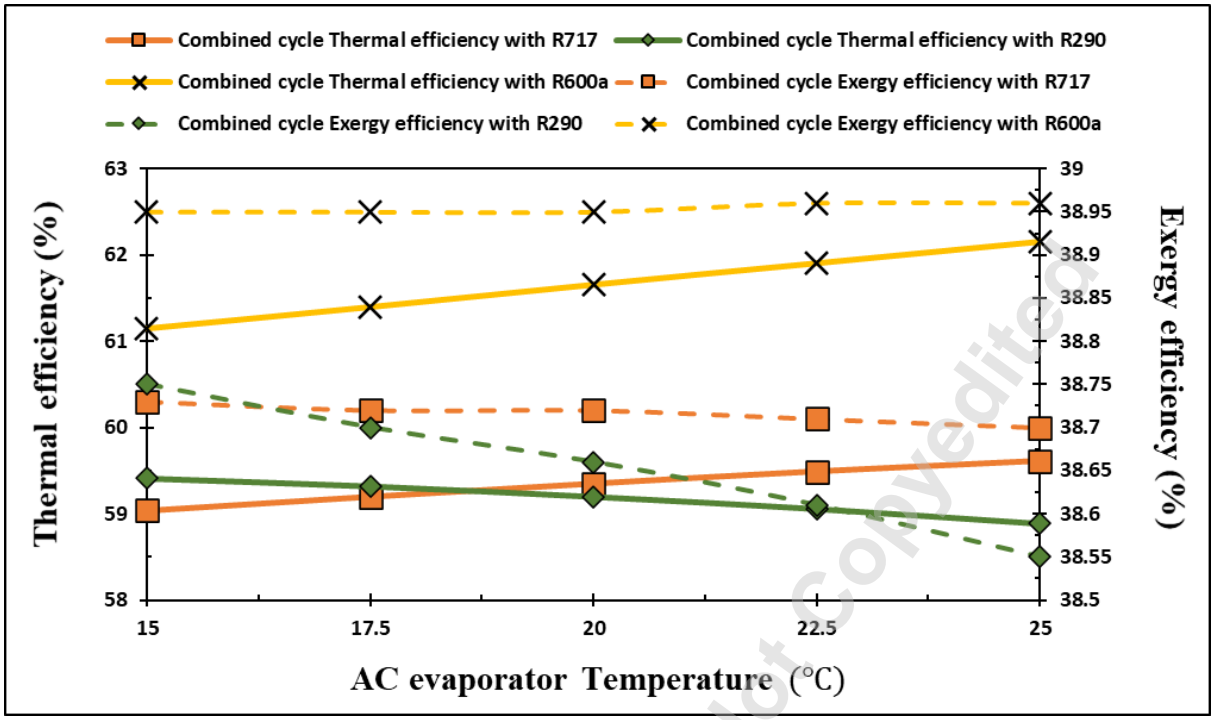


Fig. 7. Effect of AC evaporator temperature on the thermal and exergy efficiencies of combined cycle

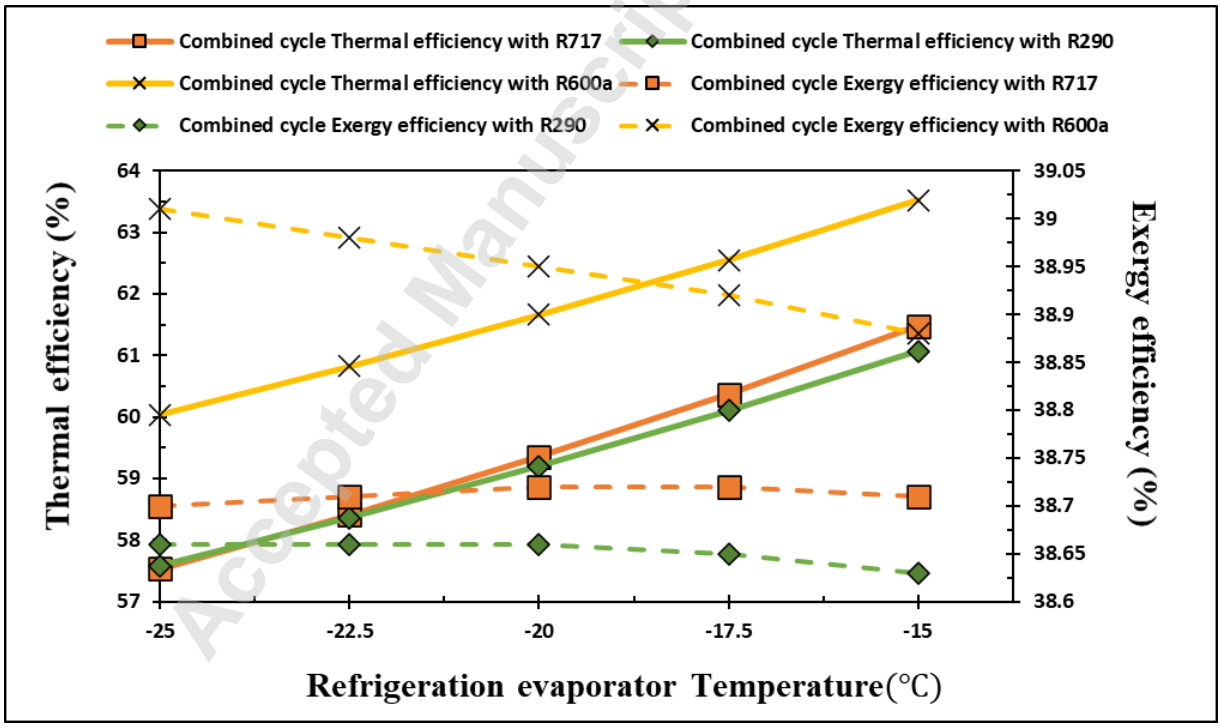


Fig. 8. Effect of refrigeration evaporator temperature on the thermal and exergy efficiencies of combined cycle

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

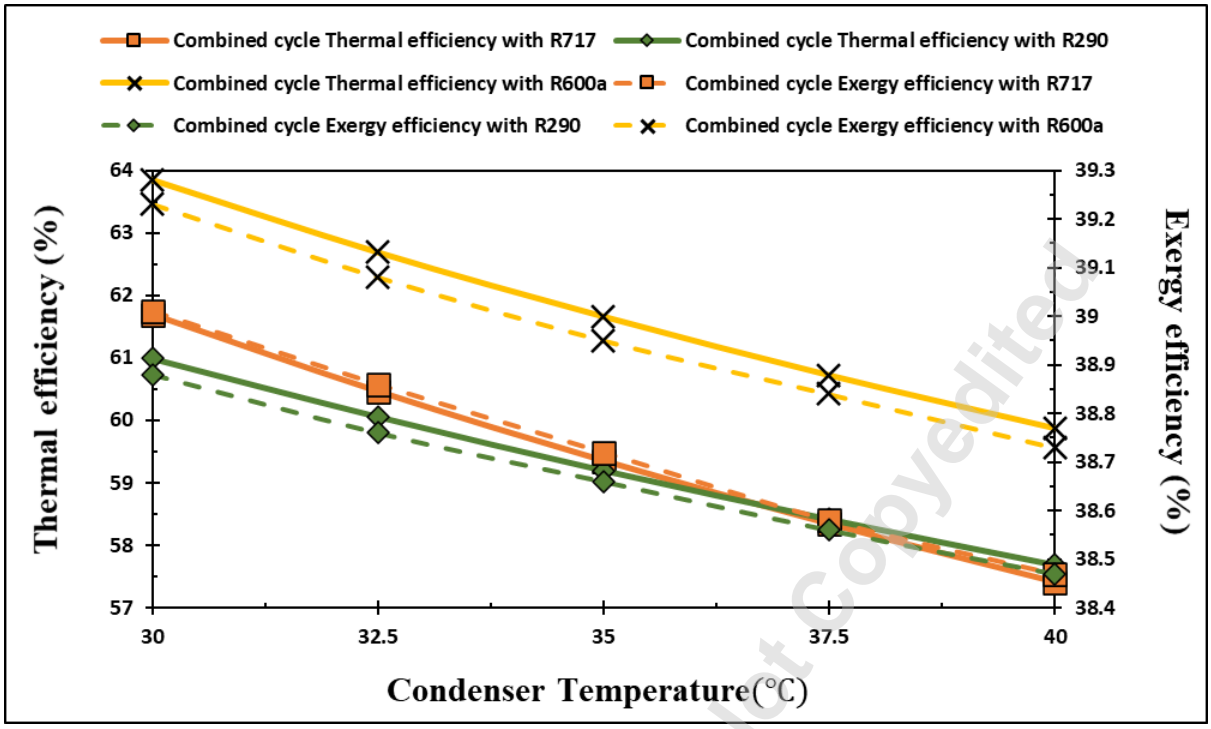


Fig. 9. Variation of thermal and exergy efficiencies of combined cycle with condenser temperature

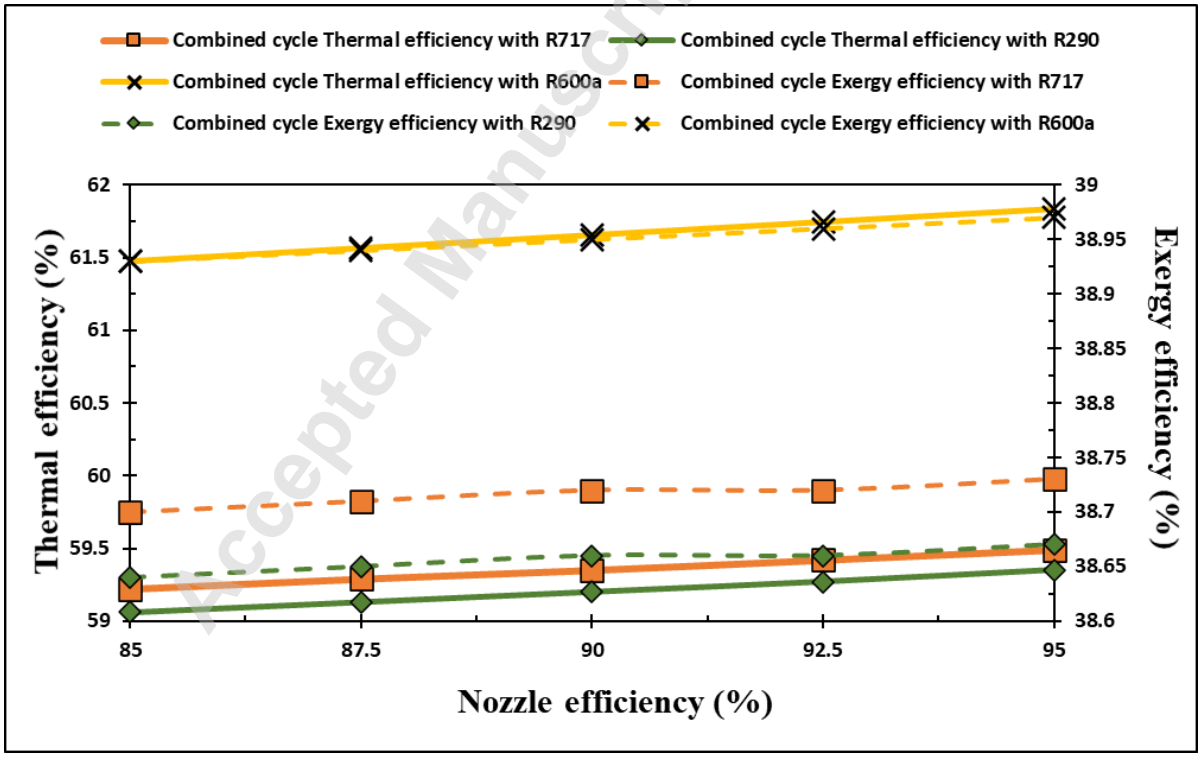


Fig. 10. Variation of thermal and exergy efficiencies of combined cycle with nozzle efficiency

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

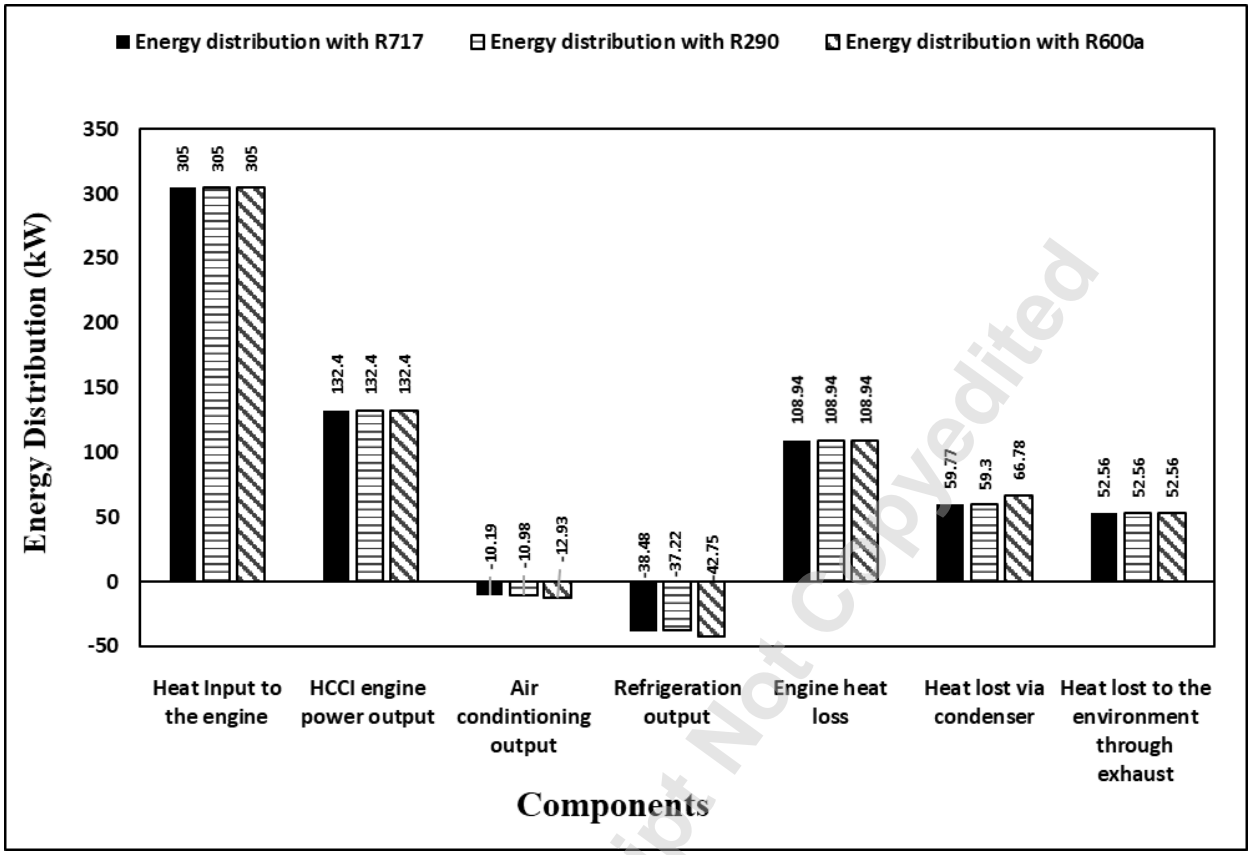


Fig. 11. Distribution of fuel energy in natural gas operated HCCI engine based combined cycle

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

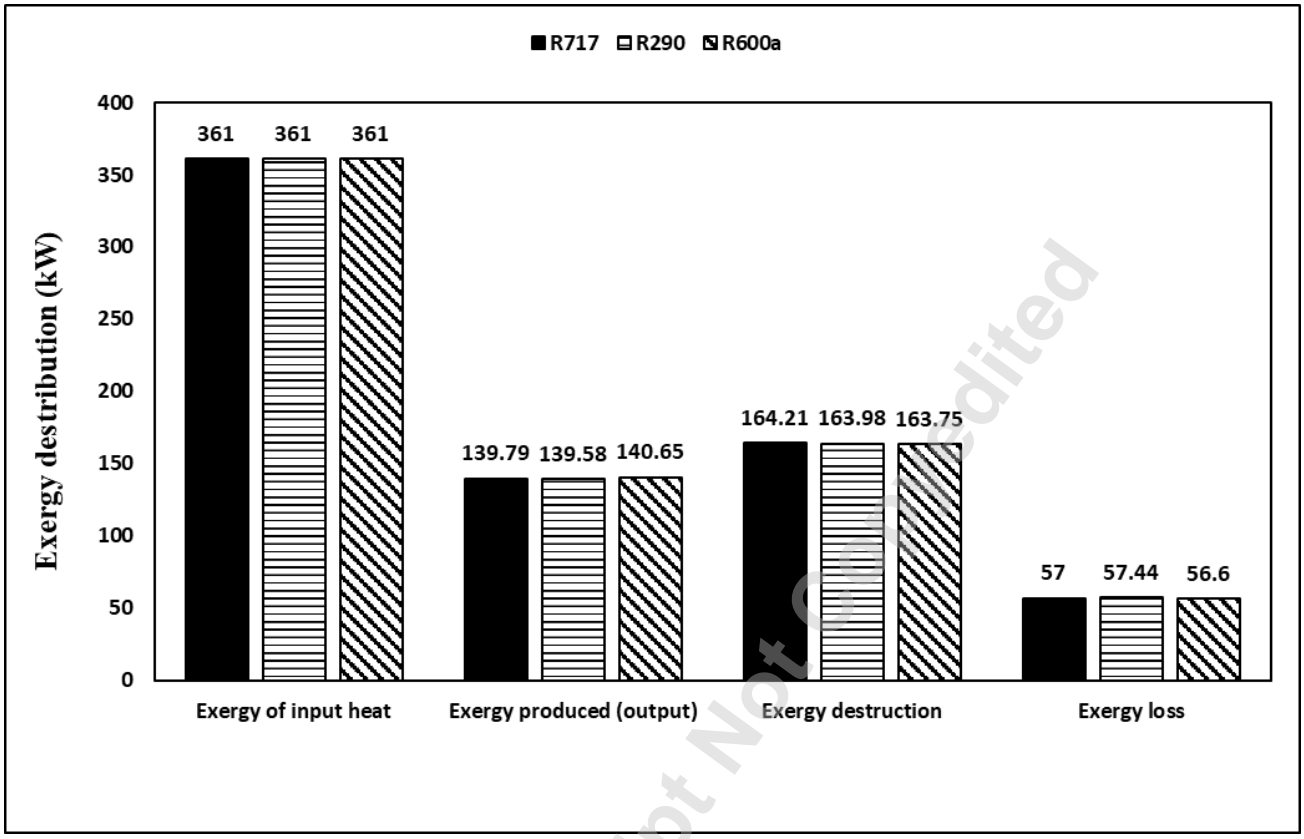


Fig. 12. Distribution of fuel exergy input in the produced, exergy destroyed and exergy loss in the proposed combined cycle

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

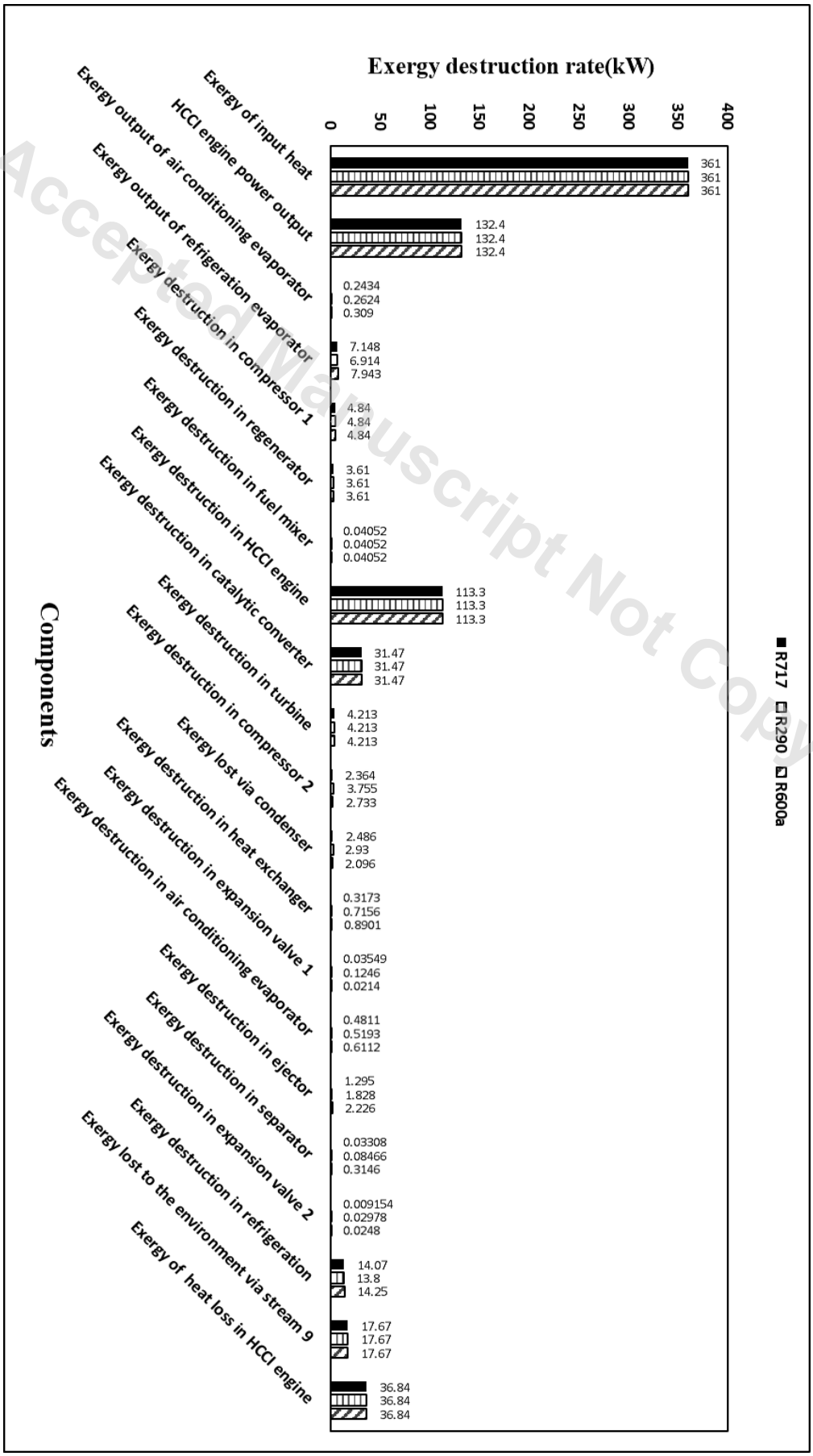


Fig. 13. Breaking down of the fuel exergy input in the components of the proposed combined cycle

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Accepted Manuscript Not Copyedited

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

Table 1: Component wise energy balance equations for natural gas operated HCCI engine based refrigeration and air-conditioning system [22, 31, 33]

Component	**Balance Equations**
Compressor (C1)	$\dot{W}_{C1} = \dot{m}_a(h_2 - h_1)$
Regenerator (Reg)	$\dot{m}_a\,(h_3 - h_2) = \dot{m}_{exh}\,(h_8 - h_9)$
Fuel mixer (FM)	$\dot{m}_a\,h_3 + \dot{m}_f h_4 = (\dot{m}_a + \dot{m}_f)h_5$
HCCI engine (Compression) process,1´- 2´	$(\dot{m}_a + \dot{m}_f + \dot{m}_{rg})(h_{2'} - h_{1'}) = \dot{W}_{comp} - \dot{Q}_{surr}$
Heat addition process, 2´- 3´	$(\dot{m}_a + \dot{m}_f + \dot{m}_{rg})(h_{3'} - h_{2'}) = \dot{Q}_H$
Expansion process,3´- 4´	$(\dot{m}_a + \dot{m}_f + \dot{m}_{rg})(h_{3'} - h_{4'}) = \dot{W}_{exp} + \dot{Q}_{surr}$
Blowdown Process, 4´-5´	$\dot{m}_{exh}\,h_{4'} = \dot{m}_{exh}\,h_{5'} + \dot{Q}_{surr}$
Catalytic convertor (CC)	$\sum_{Product} n_e\left(\bar{h}_f^{\circ} + \Delta\bar{h}\right)_e = \sum_{Reactant} n_i\left(\bar{h}_f^{\circ} + \Delta\bar{h}\right)_i$
Turbine (T)	$\dot{W}_{T1} = \dot{m}_{exh}\,(h_7 - h_8)$
Compressor (C2)	$\dot{W}_{C2} = \dot{m}_{10}(h_{10} - h_{20})$
Condenser (Cond)	$\dot{Q}_{Cond} = \dot{m}_{10}(\,h_{10} - h_{11})$
Heat exchanger (HX)	$\dot{Q}_{HX} = \dot{m}_{11}(h_{11} - h_{12}) = \dot{m}_{20}(h_{20} - h_{19})$
Expansion valve (EV1)	$h_{12} = h_{13}$
Air-conditioner (AC)	$\dot{Q}_{AC} = \dot{m}_{14}\,(h_{14} - h_{13})$
Ejector (Eje)	$\dot{m}_{14}h_{14} + \dot{m}_{18}h_{18} = \dot{m}_{15}h_{15}$
Separator (Sep)	$\dot{m}_{16}h_{16} + \dot{m}_{19}h_{19} = \dot{m}_{15}h_{15}$
Expansion valve (EV2)	$h_{16} = h_{17}$
Refrigerator (ref)	$\dot{Q}_{ref} = \dot{m}_{18}\,(h_{18} - h_{17})$

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

Table 2: Component wise exergy balance equations for natural gas operated HCCI engine based refrigeration and air-conditioning system [22, 23, 31, 32]

Component	**Balance Equations**
Compressor (C1)	$\dot{\mathrm{E}}_{\mathrm{x,D,C1}} = \dot{\mathrm{m}}_1(e_{x,1} - e_{x,2}) + \dot{W}_{C1}$
Regenerator (Reg)	$\dot{\mathrm{E}}_{\mathrm{x,D,Reg}} = \dot{\mathrm{m}}_2(e_{x,2} - e_{x,3}) + \dot{\mathrm{m}}_8\,(e_{x,8} - e_{x,9})$
Fuel mixer (FM)	$\dot{\mathrm{E}}_{\mathrm{x,D,FM}} = \dot{\mathrm{m}}_3 e_{x,3} + \dot{\mathrm{m}}_4 e_{x,4} - \dot{\mathrm{m}}_5 e_{x,5}$
HCCI engine (Compression) process,1´-2´	$\dot{\mathrm{E}}_{\mathrm{x,D,comp}} = \dot{\mathrm{m}}_5(e_{x,1'} - e_{x,2'}) + \dot{W}_{comp} - \dot{Q}_{surr}\left(1 - \frac{T_0}{T}\right)$
Heat addition process, 2´-3´	$\dot{\mathrm{E}}_{\mathrm{x,D,Heat\ addition}} = \dot{\mathrm{m}}_5(e_{x,2'} - e_{x,3'}) + \dot{Q}_H\left(1 - \frac{T_0}{T_H}\right)$
Expansion process, 3´- 4´	$\dot{\mathrm{E}}_{\mathrm{x,D,exp}} = \dot{m}_5(e_{x,3'} - e_{x,4'}) - \dot{W}_{exp} - \dot{Q}_{surr}\left(1 - \frac{T_0}{T}\right)$
Blowdown Process, 4´-5´	$\dot{\mathrm{E}}_{\mathrm{x,D},exh\ \&\ blowdown} = \dot{m}_5(e_{x,4'} - e_{x,5'}) - \dot{Q}_{surr}\left(1 - \frac{T_0}{T}\right)$
Catalytic convertor (CC)	$\dot{E}_{x,D,CC} = \dot{m}_6\left(e_{x,6}^{ph} - e_{x,7}^{ph} + e_{x,cat}^{ch}\right)$
Turbine (T)	$\dot{\mathrm{E}}_{\mathrm{x,D,T}} = \dot{\mathrm{m}}_7(e_{x,7} - e_{x,8}) - \dot{W}_T$
Compressor (C2)	$\dot{\mathrm{E}}_{\mathrm{x,D,C1}} = \dot{\mathrm{m}}_1(e_{x,20} - e_{x,10}) + \dot{W}_{C2}$
Condenser (Cond)	$\dot{\mathrm{E}}_{\mathrm{x,D,Cond}} = \dot{\mathrm{m}}_{10}(e_{x,10} - e_{x,11}) - \dot{Q}_{Cond}\left(1 - \frac{T_0}{T_{Cond}}\right)$
Heat Exchanger (HX)	$\dot{\mathrm{E}}_{\mathrm{x,D,HX}} = \dot{\mathrm{m}}_{11}(e_{x,11} - e_{x,12}) + \dot{\mathrm{m}}_{20}(e_{x,19} - e_{x,20})$
Expansion valve (EV1)	$\dot{\mathrm{E}}_{\mathrm{x,D,EV1}} = \dot{\mathrm{m}}_{12}(e_{x,12} - e_{x,13})$
Air-conditioner (AC)	$\dot{\mathrm{E}}_{\mathrm{x,D,AC}} = \dot{\mathrm{m}}_{13}(e_{x,13} - e_{x,14}) - \dot{Q}_{AC}\left(\frac{T_0}{T_{AC}} - 1\right)$
Ejector (Eje)	$\dot{\mathrm{E}}_{\mathrm{x,D,Eje}} = \dot{\mathrm{m}}_{14} e_{x,14} + \dot{\mathrm{m}}_{18} e_{x,18} - \dot{\mathrm{m}}_{15} e_{x,15}$
Separator (Sep)	$\dot{\mathrm{E}}_{\mathrm{x,D,Sep}} = \dot{\mathrm{m}}_{15} e_{x,15} - \dot{\mathrm{m}}_{16} e_{x,16} - \dot{\mathrm{m}}_{19} e_{x,19}$
Expansion valve (EV2)	$\dot{\mathrm{E}}_{\mathrm{x,D,EV2}} = \dot{\mathrm{m}}_{16}(e_{x,16} - e_{x,17})$
Refrigerator (ref)	$\dot{\mathrm{E}}_{\mathrm{x,D,ref}} = \dot{\mathrm{m}}_{17}(e_{x,17} - e_{x,18}) - \dot{Q}_{ref}\left(\frac{T_0}{T_{ref}} - 1\right)$

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

Table 3: Characteristics of Refrigerants

Refrigerant	**R717**	**R290**	**R600a**
Chemical Formula	NH_3	C_3H_8	C_4H_{10}
Normal boiling point (ºC)	-33.35	-42.2	-12
Critical pressure (kPa)	11277	4247	3640
Critical temperature (ºC)	132.4	96.7	134.7
Triple point Temperature (ºC)	-77.5	-187.1	-159.6
Molecular Weight (kg/kmol)	17.03	44.1	58.12
Safety class	B2	A3	A3
GWP (100 years)	0	<3	<3
ODP	0	0	0

Table 4: Input data used for the computation of results of the natural gas operated HCCI engine based refrigeration and air-conditioning system

Environmental temperature (T_0) (℃)	30
Environmental pressure (P_0) (kPa)	101.325
Effectiveness of regenerator	0.78
Turbocharger efficiency (%)	80
Compression ratio of engine	16:1
Catalytic converter	Oxidizing type catalyst, Pd-loaded SiO_2–Al_2O_3
Volume of displacement per cylinder	2400 cm^3
Diameter of cylinder	13.7 cm

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

Journal of Energy Resources Technology. Received May 30, 2021;
Accepted manuscript posted August 19, 2021. doi:10.1115/1.4052248
Copyright © 2021 by ASME

Stroke	16.5 cm
Connecting rod length	26.2 cm
Fraction of residual gases	0.03
Speed of engine	1400 - 2200 rpm
Engine volumetric efficiency (%)	100
Combustion process in HCCI engine	Controlling the bulk auto ignition mechanism using chemical kinetics
Equivalence ratio	0.3-0.9
Condenser Temperature (℃)	30-40
Effectiveness of heat exchanger	0.8
AC evaporator Temperature (℃)	15-25
Refrigeration evaporator Temperature(℃)	(-15) - (-25)
Nozzle efficiency (%)	85-95
Mixing chamber efficiency (%)	90
Diffuser efficiency (%)	85

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/energyresources/article-pdf/doi/10.1115/1.4052248/6742876/jert-21-1477.pdf by Delhi Technological University user on 31 August 2021

scientific reports

Check for updates

OPEN

Measuring landslide vulnerability status of Chukha, Bhutan using deep learning algorithms

Sunil Saha[1], Raju Sarkar[2]✉, Jagabandhu Roy[1], Tusar Kanti Hembram[3], Saroj Acharya[4], Gautam Thapa[5] & Dowchu Drukpa[6]

Landslides are major natural hazards that have a wide impact on human life, property, and natural environment. This study is intended to provide an improved framework for the assessment of landslide vulnerability mapping (LVM) in Chukha Dzongkhags (district) of Bhutan. Both physical (22 nos.) and social (9 nos.) conditioning factors were considered to model vulnerability using deep learning neural network (DLNN), artificial neural network (ANN) and convolution neural network (CNN) approaches. Selection of the factors was conceded by the collinearity test and information gain ratio. Using Google Earth images, official data, and field inquiry a total of 350 (present and historical) landslides were recorded and training and validation sets were prepared following the 70:30 ratio. Nine LVMs were produced i.e. a landslide susceptibility (LS), one social vulnerability (SV) and a relative vulnerability (RLV) map for each model. The performance of the models was evaluated by area under curve (AUC) of receiver operating characteristics (ROC), relative landslide density index (R-index) and different statistical measures. The combined vulnerability map of social and physical factors using CNN (CNN-RLV) had the highest goodness-of-fit and excellent performance (AUC = 0.921, 0.928) followed by DLNN and ANN models. This approach of combined physical and social factors create an appropriate and more accurate LVM that may—support landslide prediction and management.

Among the numerous natural hazards, landslides are considered to be one of the biggest, as they can cause tremendous loss of life and property as well as affect the natural ecosystem. Now attention is being shifted towards the issue of landslides, as increasing developmental works in the Himalayan areas are taking place[1]. The Himalayan region's tectonic fragility is already illustrated, with an overabundance of literature focusing on the broad scale and few on the micro-scale. Being a part of this region, Bhutan has long been known as a place sensitive to natural disasters, such as landslides[2]. Therefore, the need for micro-scale study i.e. province-wise study in Bhutan can be fruitful to act locally viz. the present analysis.

In recent decades, the most suitable way to tackle this hazardous event is the spatial assessment of vulnerability to landslides[3,4]. Assessment of geomorphological, geological, tectonic, climate, vegetation, and land practices may help in identifying the area susceptible and vulnerable to landslides[5,6]. The occurrence of landslides is a natural phenomenon but man-made processes are one of the causes of vulnerability to landslides, which makes it difficult to predict the spatial and temporal occurrence of landslides[7]. In this connection, landslide vulnerability maps (LVMs) may be used as the primary method to classify the high-risk zones which are prone to landslides and also help in identifying the variables responsible for the occurrence of landslides[8]. Presently, with the development of various tools, hardware, and availability of data, it has become easy to produce landslide susceptibility and vulnerability maps[9].

In spatial landslide modelling, the terms "susceptibility" and "vulnerability" are often used interchangeably; however, "susceptibility" often points to causes intrinsic to physical predisposition (e.g., structural and topographical), while "vulnerability" often corresponds to external influences along with causes intrinsic to physical predisposition (e.g., anthropogenic exposure)[10]. Most of the previous landslide studies have not considered the "vulnerability" aspect rather those have modelled the "susceptibility". Therefore, another exclusivity of the present work also lies in it. According to the report (May 2010) of XVIth SAARC (The South Asian Association of Regional Cooperation) summit hosted by the Royal Government of Bhutan, several social factors are also

[1]Department of Geography, University of Gour Banga, Malda, West Bengal, India. [2]Department of Civil Engineering, Delhi Technological University, Delhi, India. [3]Department of Geography, Nistarini College, Purulia, West Bengal, India. [4]National Centre for Hydrology and Meteorology, Thimphu, Bhutan. [5]Royal Government of Bhutan Projects, Phuentsholing, Bhutan. [6]Department of Geology and Mines, Ministry of Economic Affairs, Thimphu, Bhutan. ✉email: rajusarkar@dce.ac.in

responsible for landslides. Among them, landslides caused by road cutting are very common. Landslides are more likely due to intensive deforestation for farmland. The potential human influence could well be the outcome of building roadways and other civil constructions on the slopes which are devoid of planning. Blasting is another addendum to these. The landslide problem is only going to get worse as the world's population grows, putting more strain on natural resources. Therefore, the inclusion of these man-made processes as causative factors to produce a spatial vulnerability map is of utmost importance.

Several studies in the past have considered the role of the geo-environmental factors while few have looked at the socio-economic aspects of landslide events. Based on the arrangement of anthropogenic and physical elements in a region, landslides can be accelerated by different geo-environmental as well as socio-economic factors. With the increasing requirements of the growing population, modification of the land cover for agricultural expansion, road and civil construction, and tourism has influenced the stability of slopes. Therefore, an effort has been made to analyze the role of both physical and social aspects for spatial vulnerability mapping of landslides in this study[1].

Integration of RS-GIS techniques with knowledge-driven methods or data-mining methods in the landslide susceptibility as well as vulnerability assessment using spatial and non-spatial data has been widely used[11.] The knowledge-driven multi-criteria decision approaches (MCDM) and machine learning models (ML) include frequency ratio, evidential belief function, logistic regression, weight-of-evidence, fuzzy logic, artificial neural network, support vector machine, random forest, logistic model tree, boosted regression tree, etc., which are very widely used in modelling the landslide susceptibility rather than landslide vulnerability[11,12]. In most cases, the ML models provided superior results compared to the traditional approaches because the non-linear data can be adequately treated by the ML models with various scales[13]. Numerous works have reported that the ensemble of ML and conventional statistical models provided superior results compared to the single ML model for modelling landslides[14,15].

Along with ML models, recent works have acknowledged deep learning models i.e. convolution neural network (CNN), deep learning neural network (DLNN), recurrent neural network (RNN), etc. as emergent and more powerful tools for spatial modelling because of better results than the conventional ML models[16,17.] DLNN shows many topologies as they appear with more than one hidden layer than that of the single-hidden-layer neural network. It is used for the extraction, transformation, pattern recognition, and classification tasks of supervised or unsupervised features[17]. In the case of CNN model, high detection accuracy for landslide detection was achieved by Yu et al.[18]. Ghorbanzadeh et al.[19] on the other hand reported contrasting findings of the CNN approach with three futuristic methods of MLs despite feasibility and high precision. ANN model outperformed different statistical and ML models in identifying landslide-prone areas as reported by Pourghasemi and Rahmati[20]. The deep learning models used in the present study are still not used for vulnerability studies by the researchers. Previously published landslides susceptibility studies on different parts of Bhutan have revealed that none of them applied deep learning models to predict susceptibility or vulnerability.

Based on the above-reviewed literature, in the present study, ANN, CNN, and DLNN have been selected to analyze the landslide vulnerability status of the Chukha district in Bhutan. The risk posed by landslides in the study region, which is located in the southern part of the country, is mainly due to heavy rainfall during monsoon[21] and weak geology predominantly made of highly fractured and weathered phyllites, slates, and schists with a high quantity of clay minerals[22]. The situation is further aggravated due to anthropogenic activities like road cuttings[23]. The performance of the above said models were evaluated by the most selective validation methods and statistical measures. The pixels-based evaluation of landslide models has been done by the relative landslide density index (R-index) method.

Materials and methods

Study area. The Chukha Dzongkhag (district) lies in the southern foothills of the Bhutan Himalayas (Fig. 1). It shares border with the neighbouring Indian state of West Bengal and thus serves as a vital route for mutual trading between the two countries. The major border town of Phuentsholing is the gateway city that connects India to Western Bhutan. Some of the important hydropower plants such as Chukha and Tala are located in the Chukha district. Because of proximity to India and ease of accessibility to large Indian market, many of the country's industrial infrastructures are located along the foothill flat areas of Chukha district. From the statistic report of Bhutan Govt., an estimated 4.22% of the entire district (1879.77 km^2) comprises of fields cultivated with oranges and potatoes that constitutes the main source of rural cash income.

In the study area, the monsoon season extends from June–September of each year and about 78% of the annual rainfall occurs during this period, making the area highly vulnerable to landslide hazards and flooding. In addition, the rugged mountainous terrain with steep slopes further exacerbates the condition that triggers the landslides, exposing the area to the vagaries of weather events. The landslides obstruct the national highway (Phuentsholing-Thimphu Highway) which is the main trade route of the country, causing heavy losses to lives and infrastructures (Table 1). One of the major landslides in the region was along the Phuentsholing-Thimphu highway (26.85 N, 89.33 E to 27.15 N, 89.55 E) just after the 2016 monsoon. The intensity and frequency of such events are expected to increase with climate change, resulting in an increasing risk to the residents and also affecting the economy of the country.

Methodology. The steps followed in this work are presented in Fig. 2.

(1) Prepared landslide inventory map based on landslide locations identified through field investigation.
(2) Collected data for several geo-environmental and socio-economic factors.
(3) Selected factors using collinearity test and information gain ratio (IGR) methods.

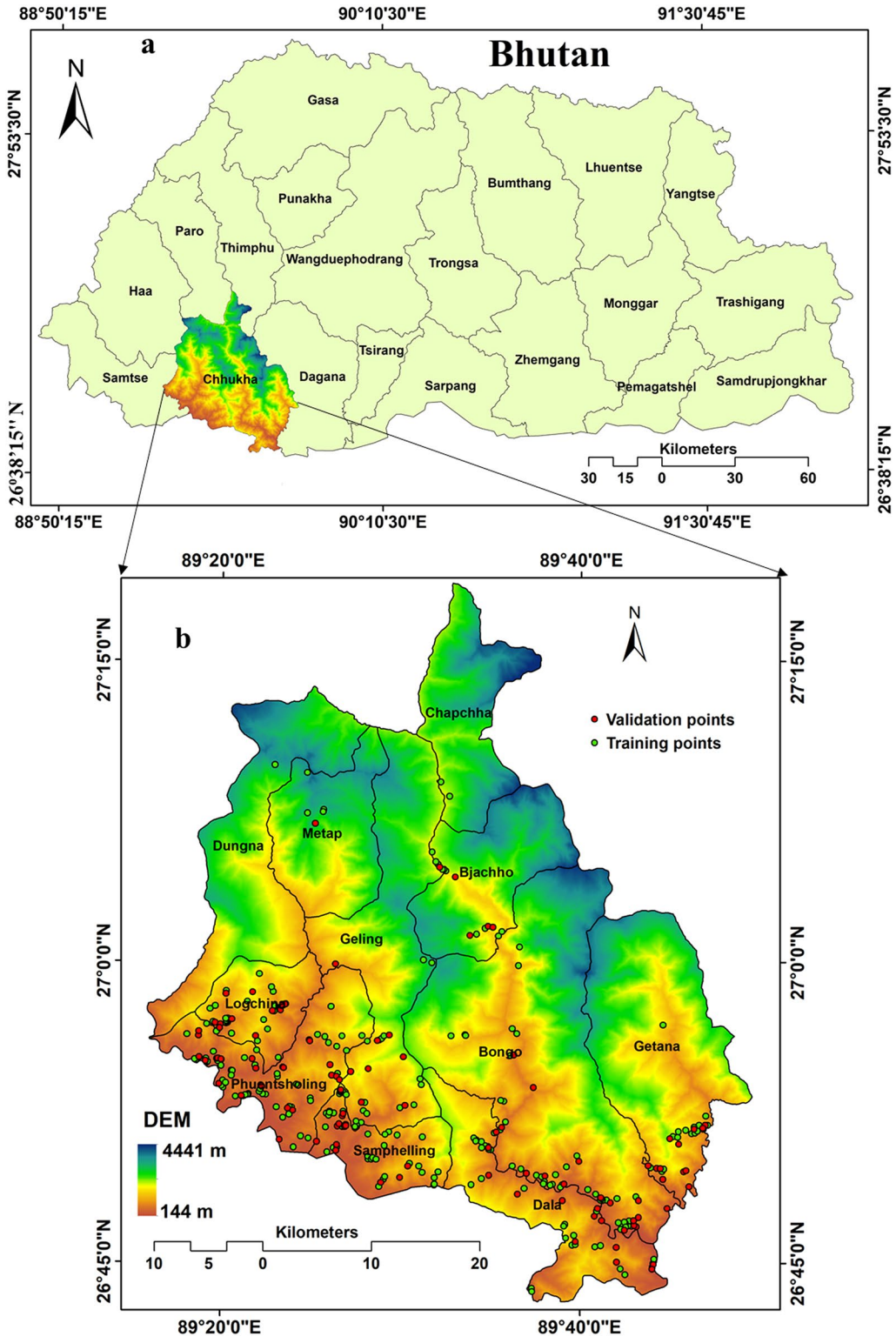


Figure 1. Location of the study area: (**a**) Bhutan and (**b**) Chukha district which is prepared by open source QGIS 3.16 software (https://qgis.org/en/site/forusers/download.html).

(4) Applied two deep learning approaches i.e. DLNN and CNN, and one benchmark machine learning model i.e. ANN to produce physical landslide vulnerability, socio-economic landslide vulnerability and overall landslide vulnerability.

Name of the area	Latitude and longitude	Date and time of landslides	Losses caused by the landslides (properties and lives)	Source
Geling gewog (Kamji)	Approx 26.90242 N 89.51761 E	June 16, 2020	Road block	Bhutan Today (National news paper, June 17, 2020)
Sorchen & Kamji	Approx (26.89472 N 89.43556 E)—Sorchen Approx 26.90242 N 89.51761 E—Kamji	July 12, 2019	Road blocks- nearly 200 vehicles stranded	BBS (National TV, July 12, 2019)
Phuentsholing gewog (Kabeytar)	Approx 26.86333 N 89.39611 E	June 25, 2019	Numerous vehicle and houses were damaged	Bhutan 24×7 (National newspaper, June 26, 2019)
Phuentsholing gewog (Darjaygang)	Approx 26.8775 N 89.38556 E	August 12, 2017	Killed a 60 year old women and cut off road connection between Dungna and Metedkha gewog	Kuensel (National newspaper, August 13, 2017)
Geling gewog (Kamji)	Approx 26.90242 N 89.51761 E	June 23, 2016	Massive landslide washed away the section of road at Kamji	Kuensel (National newspaper, June 24, 2016)

Table 1. Details of some recent major landslides and losses occurred in the study area.

(5) Analysed significant factors using RF and Chi-square attribute evaluation (CSAE) ML models.
(6) Finally, model performances were analyzed using ROC, AUC, efficiency, accuracy, true positive rate (TRP), false positive rate (FPR), true negative rate (TNR), false negative rate (FNR), Kappa index, root mean square error (RMSE), mean absolute error (MAE) and relative landslide density index (R-index) methods.

Preparation of landslide inventory map (LIM). Mapping of the past and present records of landslide events is called an inventory map[24]. In this study, these records were collected from the Project Dantak, Border Road Organisation (BRO), Govt. of India, and from two Royal Government of Bhutan organizations—Department of Geology and Mines (DGM) and National Centre for Hydrology and Meteorology (NCHM). For location verification, field survey was carried out using Global Positioning System devices in January 2020 (Fig. 3). A total of 350 landslides were mapped for modelling the relative landslide vulnerability, of which 50.4% are rock falls, 40.2% are debris slides and 9.4% are rotational slides. Correspondingly, the same amount of non-landslide points was randomly selected for training and validating the models. Out of the total landslide location, 70% was used as training dataset and the remaining 30% as testing dataset[24].

Preparation of the landslide vulnerability conditioning factors (LVCFs). In this study, 22 geo-environment and 9 socio-economic factors were used for landslide vulnerability modelling which were obtained from the high-resolution (12.5 × 12.5 m) PALSAR DEM of Alaska DEM facility, Landsat 8OLI/TIRS of 30 m × 30 m resolution from USGS Earth Explorer, the demographic data from the National Statistics Bureau, Bhutan, rainfall data of last five years from National Centre for Hydrology and Meteorology, and geological map of 1:500,000 scale from Department of Geology and Mines, Royal Government of Bhutan. The soil resistivity and textural classes (clay, silt and sand) data were extracted from laboratory tests of collected samples from the study area. Data processing and modelling were performed using software such as SPSS, MS-Excel, QGIS 3.16, and R studio.

Geo-environmental factors. The elevations of the study area vary from a maximum of 4411 m to as low as 98 m (Fig. 4a) and the slope ranges from 0° to 80° (Fig. 4b). Aspect of slope is classified into nine categories namely, Flat, North, South, West, East, Northeast, Southeast, Northwest, Southwest (Fig. 4c). The spatial distribution of plan curvature and profile curvature ranges from − 32.89 to 32.12 (Fig. 4d) and from − 45.17 to 43.52 (Fig. 4e). The spatial value of convergence index (CI) ranges from − 92.25 to 89.53 (Fig. 4f) and topographical position index varies between − 17.59 and 26.35 (Fig. 4g). The value of terrain ruggedness ranges from 0 to 40 (Fig. 4h). The spatial value of topographical wetness index varies between 0.68 and 25.41 (Fig. 4i) and valley depth varies from 0 to 817 (Fig. 4j). The length of slope and value of relative slope position of this study area ranges between 0 and 119 (Fig. 4k) and from 0 to 1 (Fig. 4l), respectively. Rainfall map was prepared based on the last 5 years' average annual precipitation data of the different stations using the Inverse Distance Weighted interpolation method. The average maximum and minimum precipitations for this study area are 4130 mm and 1357 mm, respectively (Fig. 4n). The land use/land cover (LU/LC) map of the district has been prepared from Landsat 8 OLI/TIRS satellite imagery following the supervised maximum likelihood classification method. The LU/LC categories are broad leave forest, build-up areas, mixed forest, grassland, miscellaneous, conifer forest, and agriculture respectively (Fig. 4p). The normalized differential vegetation index (NDVI) was also generated using RED and NIR band of Landsat 8OLI/TIRS imagery and values differs from − 0.33 to 0.81 (Fig. 4q). The positive values of NDVI indicate the dense vegetation cover areas, while the negative values indicate low vegetation cover areas. The geology of the study can be categorized into Buxa Group, Daling-Shumar Group, the Lesser Himalayan zone, Paro formation, and the structurally lower Greater Himalayan zone (Fig. 4r). The soil resistivity values were determined by surface electrical resistivity method using soil samples and the map depicting resistivity

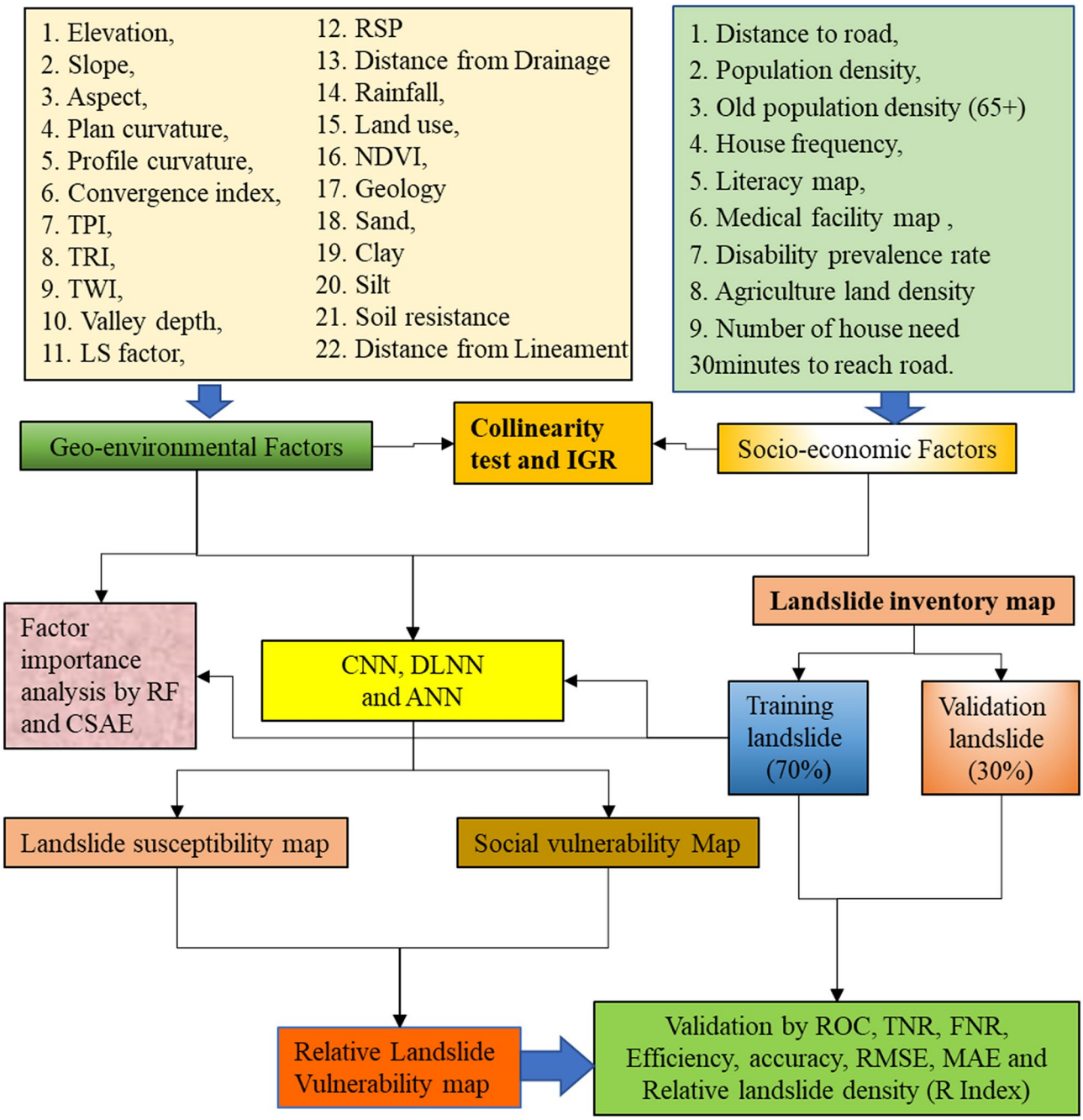


Figure 2. Flowchart of the developed methodology.

distribution was prepared applying Inverse Distance Weighted (IDW) interpolation method. The soil resistivity spatially varies from 276 to 799 Ω-m (Fig. 4v). Following the same method sand, clay, and silt maps were also prepared. In different parts of the district the percentage of sand, silt, and clay was found to be 34–50%, 12–30%, and 28–45%, respectively (Fig. 4s, 4t, 4u). Drainage of the study area has been extracted from the open series topographical maps and DEM. Buffering tool of QGIS 3.16 was used to prepare the distance from the river map (Fig. 4m) and distance from lineament map (Fig. 4o).

Socio-economic factors. To identify the social vulnerability of landslides, socio-economic factors were chosen. The distance from road was prepared using the Euclidian distance buffering tool (Fig. 5f). The demographic data was derived from the district headquarter of Chukha district, Bhutan. The thematic maps of the population density, old population density, house frequency, literacy rate, medical facility, disability prevalence rate, number of households that require at least 30 min to reach the road head (Fig. 5a–i) were prepared in GIS platform based on the block-wise data.



Figure 3. Some field photos of landslide events in Sharphug (latitude 26.76 N, longitude 89.7 E), under Darla Gewog (sub-division) of Chukha Dzongkhag (District).

Factor selection techniques. Collinearity analysis and information gain ratio was used for selecting the appropriate factors to measure the landslide vulnerability of the Chukha district.

Collinearity analysis (CA). Collinearity is a linear association between two explanatory factors. Two factors are perfectly collinear if there is an exact linear relationship between them which can influence the model result. The collinearity test was performed using the Variance Inflation Factor (VIF) and Tolerance (TOI)[25]. A tolerance of less than 0.2 and a VIF of > 5 indicates a multicollinearity problem[26]. Based on the diagnosis, a total of 22 geo-environment and 9 socio-economic factors have been selected to proceed with the modelling.

Information gain ratio (IGR) method. The information gain ratio (IGR) is one of the machine learning techniques[27] which evaluates the correlations of landslide occurrence with landslide conditioning factors and the role of these factors in the frequency of their correlations[28]. The IGR is employed to reduce a bias to multi-value property, taking into account the size and number of sections, when selecting a function. The IGR for element TWI, for example, is estimated as:

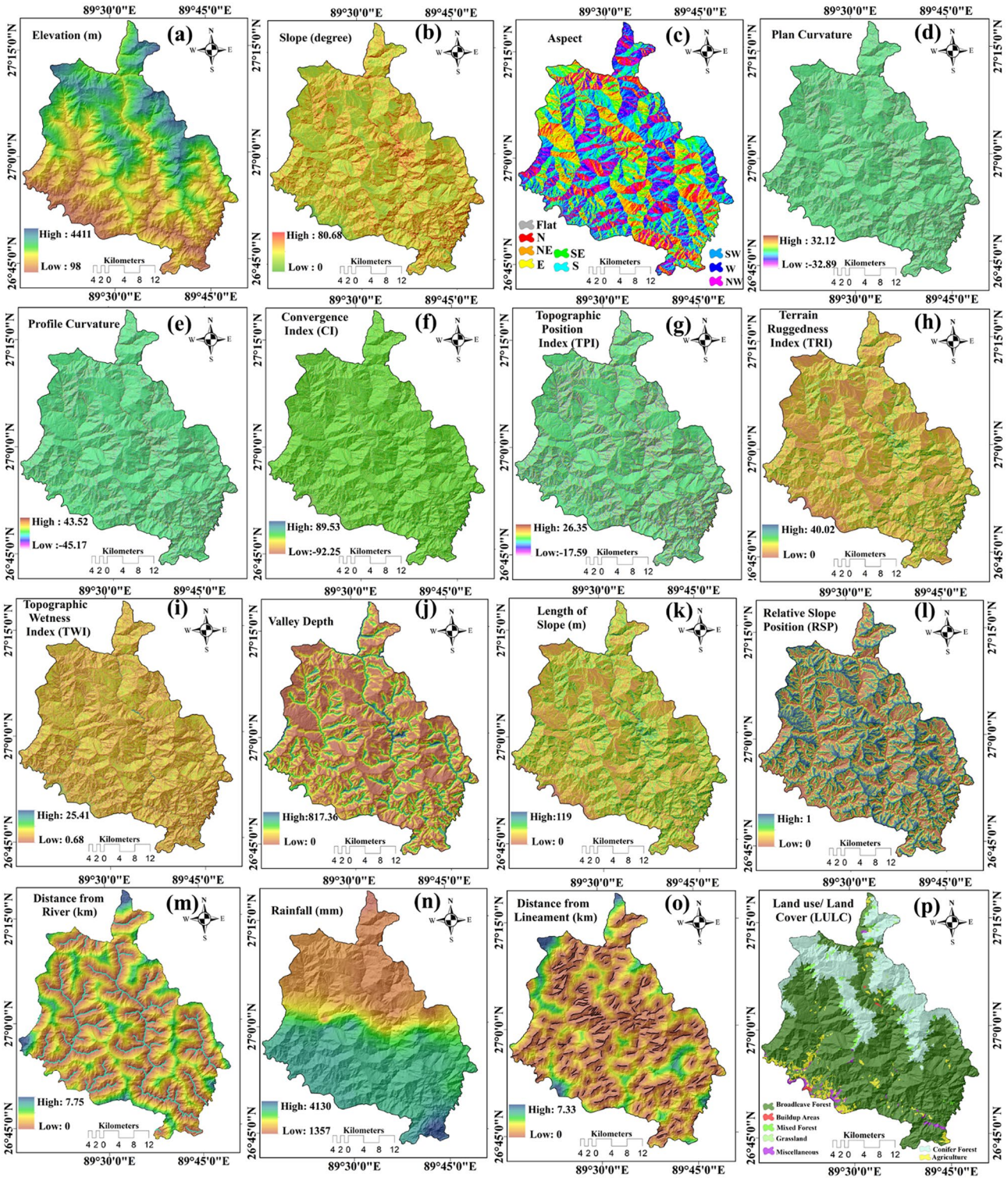


Figure 4. Factors used for producing the landslide susceptibility maps—(**a**) Elevation, (**b**) Slope, (**c**) Aspect, (**d**) Plan curvature, (**e**) Profile curvature, (**f**) Convergence index, (**g**) Topographical position index, (**h**) Terrain ruggedness index, (**i**) Topographical wetness index, (**j**) Valley depth, (**k**) Length of slope, (**l**) Relative slope position, (**m**) Distance from river, (**n**) Rainfall, (**o**) Distance from lineament, (**p**) Land use/land cover, (**q**) Normalized differential vegetation index (NDVI), (**s**) Geology map, (**s**) Sand, (**t**) Clay, (**u**) Silt, (**v**) Soil resistivity.

$$IGR(S, TWI) = \frac{Info(S) - Info(S, TWI)}{SplitInfo(S, TWI)} \quad (1)$$

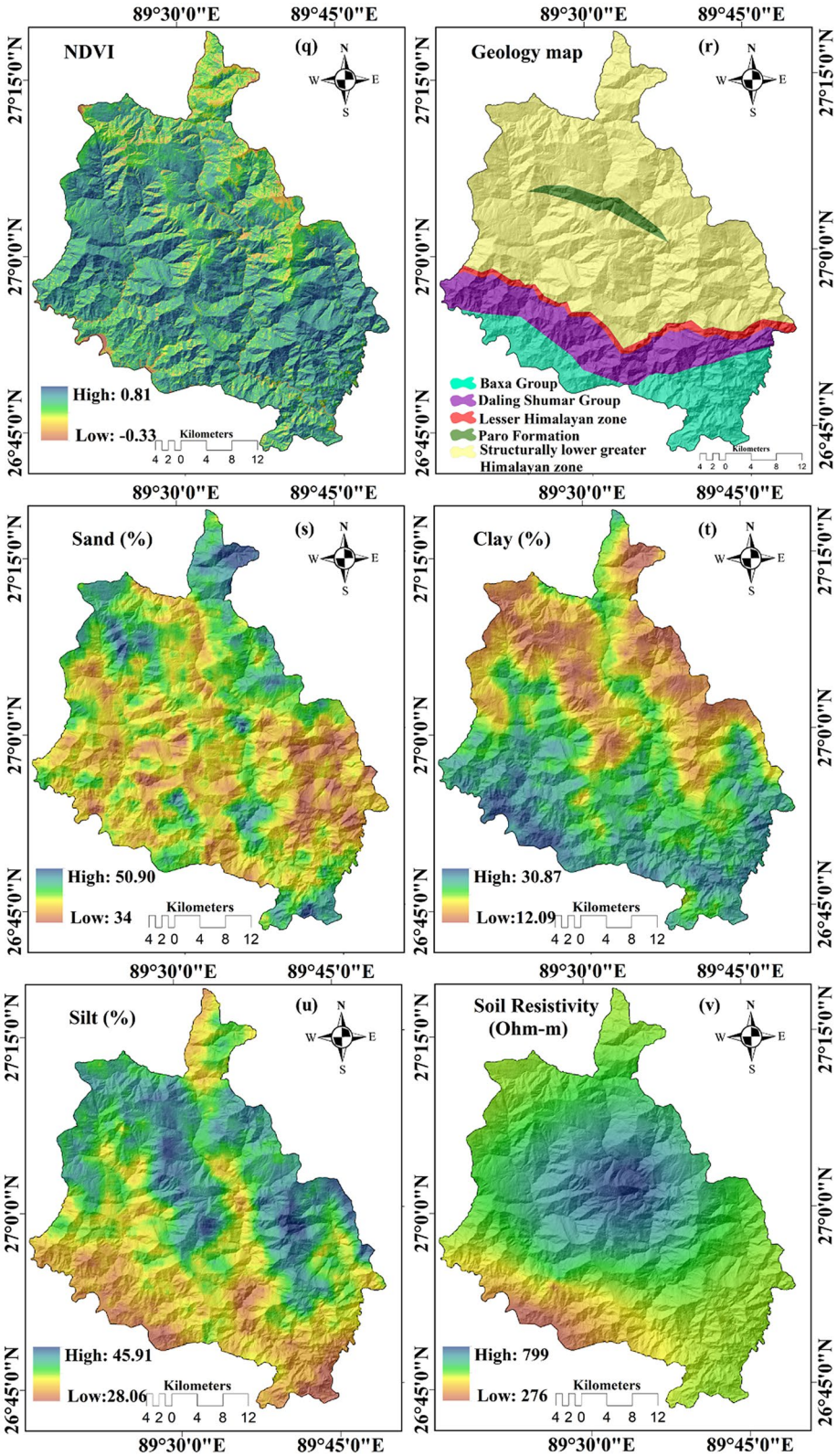


Figure 4. (continued)

$$SplitInfo(S, TWI) = -\sum_{j=1}^{m} \frac{|S_j|}{|S|} \log_2 \frac{|S_j|}{|S|} \quad (2)$$

where SplitInfo and S are represented potential information conducted by partitioning S into m subset and training[28].

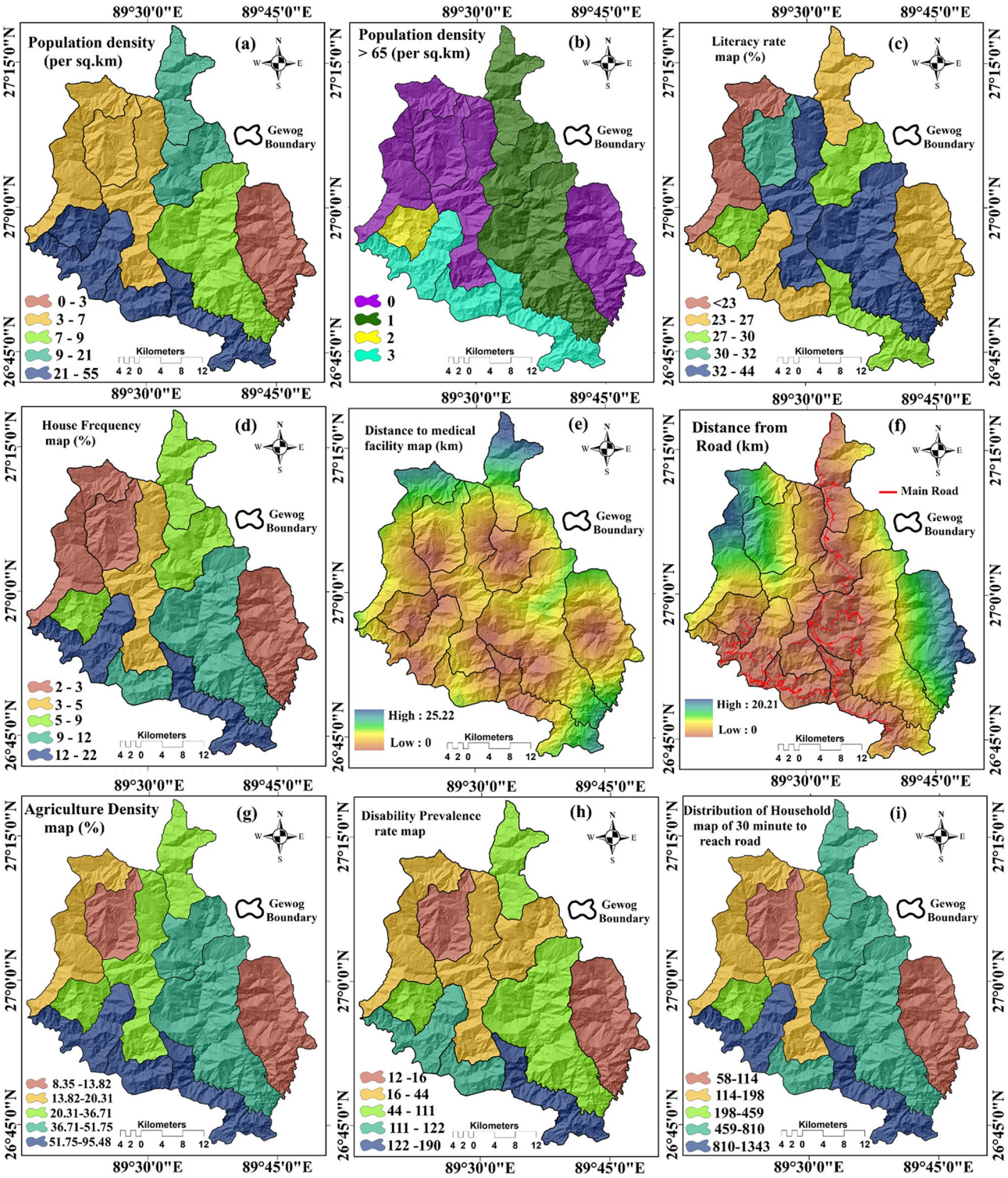


Figure 5. LVCFs used for producing the socio-economic landslide vulnerability maps: (**a**) Population density, (**b**) Old population density, (**c**) Literacy rate, (**d**) House frequency, (**e**) Distance to medical facility, (**f**) Distance to road, (**g**) Agriculture density, (**h**) Disability prevalence rate, (**i**) Household need 30 min to reach road.

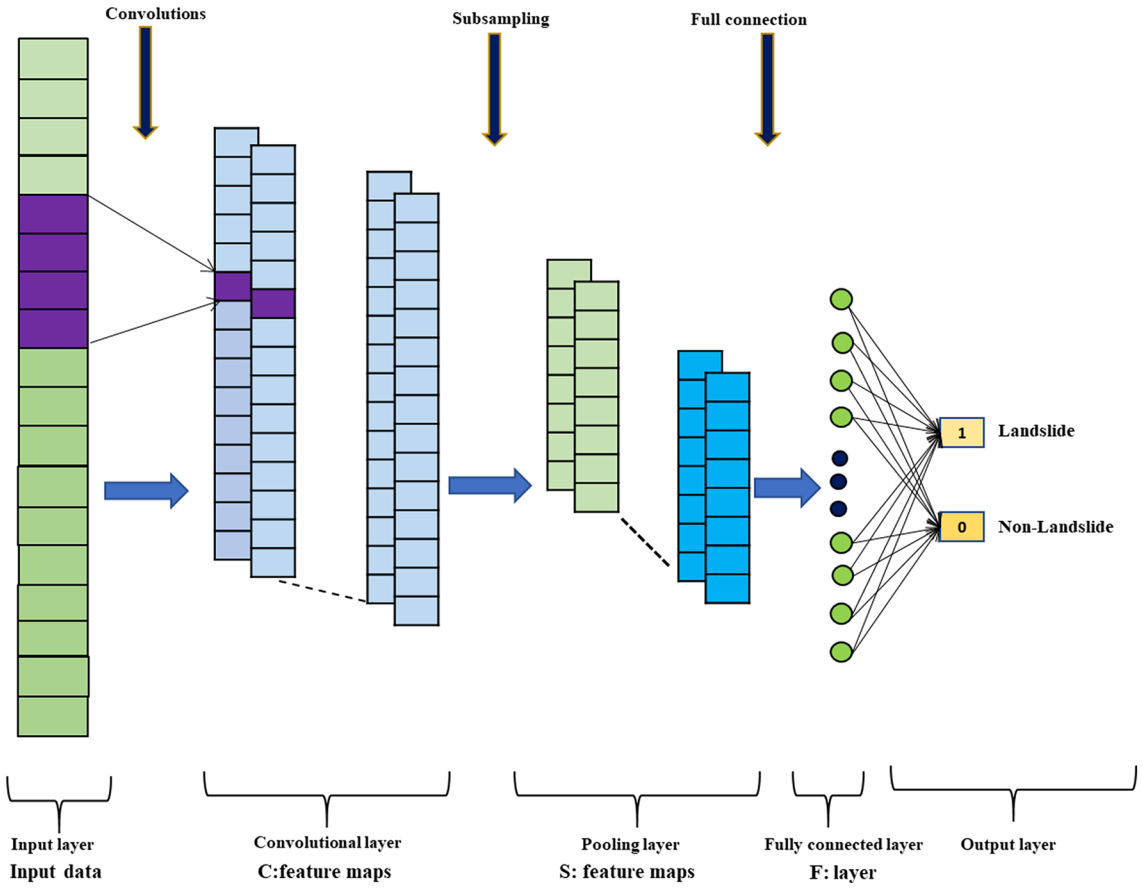


Figure 6. Theoretical structure of used CNN model.

Applied deep learning and benchmark machine learning approaches

Artificial neural network (ANN). One of the most popular ANN used for landslide susceptibility is the Multi-layer perceptron Neural Network (MLP-NN)[29]. Basically, three layers which include an input layer, one hidden layer, and an output layer were involved in the topology of MLP-NN models[30]. Although their output is regulated by their structure, the activation functions and the updating of the link weights between the processing components[31]. The MLP-NN was introduced in the current study based on the "RSNNS" R package[32] and by using a grid search technique, the number of the hidden processing elements was tuned.

Convolution neural network (CNN). DLAs are modelled on the structure of the human brain and based on an ANN. CNN, introduced by LeCun et al.[33] is a well-known deep learning algorithms (DLA). Recently, a number of disciplines, including earth science, have been increasingly using CNN for classification and prediction[34]. The utilisation of several layers, pooling, local connections, and mutual weighting distinguishes CNN from traditional neural networks. The central concept behind CNN is that images are used as input parameters. The set of indicators can be greatly decreased and processing can be accelerated. The convolution layers (CLs), pooling layers (PLs), and linear rectified unit layers are involved in the typical structure of CNN model (Fig. 6). CLs deliver the best data classification results by learning the convolutions. By decreasing the quantity of convolution architectures, PLs control overfitting that allows consistent conversion and enhances computational efficiency[35]. The ReLU improves the network's nonlinear capabilities by ReLU activation. Researchers have created and applied numerous analysis structures based on data form, image, and purpose, including GoogleNet[36], ZFNet[37], and VGGNet[38] among the more common structures. Several articles have clarified these layer forms, their basic learning criteria, and how the CNN processes the data for training[33]. The CNN-2D structure is used here because it is relevant for earth science studies[39]. The input data in CNN must be 1-D images: a 1-D input grid cell (vector) with different features which be translated into a 2-D input grid cell to ensure optimum initialization efficiency and this technique is used for mapping landslide vulnerability. In current research, we compare the number of LCVFs to the number of attribute values for each factor, and the larger of the two numbers is used to determine the size of the related two-dimensional matrix. The research region, for example, contains 25 geological categories, which is more than the number of LCVFs. As a result, for each grid cell, we created a 25×25 matrix. Since there is no sorting of data and analysis is continuous, the images are large.

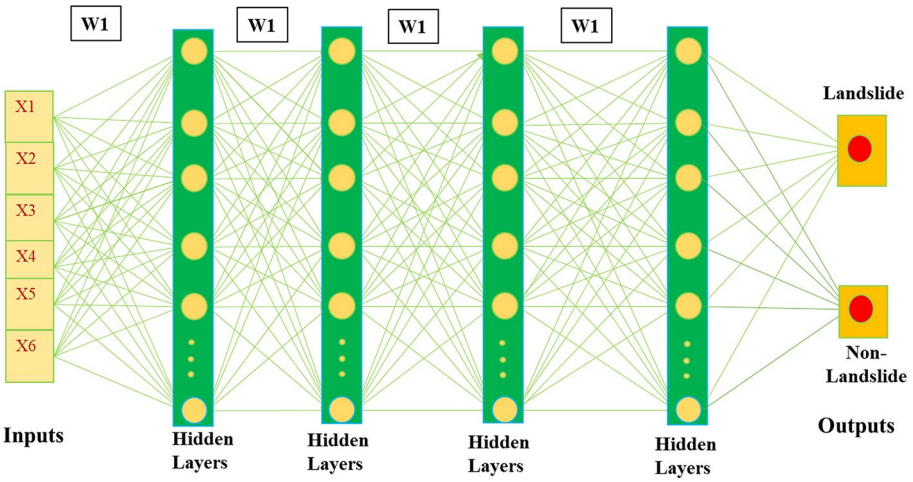


Figure 7. Configuration of the DLNN approach.

Deep learning neural networks (DLNN). The major advantage of DLNN model is that it uses raw dataset to create a high-level function[18]. DLNN comprises of three layers—an input layer, followed by hidden layers which lead to output layers[40]. Figure 7 demonstrates the conceptual setup of the DLNN model included in this analysis. The overall pattern of DLNN model is to work in such a way that the input layer delivers the signals that are diverse landslide factors, and after processing and interpretation of this information in multiple hidden layers, the impacts are shown in the model's last year, the output layer. The output layer has two probable labels, i.e. a negative label (non-landslides) and a positive label (landslides). From the last hidden layer, these classification results are collected and displayed in the output layer[41]. DLNN has unique advantages over the conventional ML algorithm, and therefore much more attention has been put on the use of the DLNN model in the area of prediction analysis. DLNN outperforms all other ML models in several ways, by making optimum use of unstructured data by specific observations to recognize the training dataset, being versatile enough as to identify new data, and being able to create new learning models by introducing more layers to the neural network architecture.

Mathematical equation mentioned below was applied in DLNN as per the Kim[40]:

$$h(x) = \begin{cases} x \ if \ x > 0 \\ 0 \ if \ x < 0 \end{cases} = \max(0, x) \ h(x) = \begin{cases} x \ if \ x > 0 \\ 0 \ if \ x \leq 0 \end{cases} = \max(0, x) \tag{3}$$

where x is an input signal, and h is an activation function. The following can be represented on the basis of the ReLU activation function as:

$$h(x) = \begin{cases} 1 \ if \ x > 0 \\ 0 \ if \ x \leq 0 \end{cases} \tag{4}$$

The cost function is the distinction between class outcomes that are experiential and expected. The loss function (L) of a cross-entropy is used to detect patterns and is given by:

$$L = -\frac{1}{N_D} \sum_{n=1}^{N_D} T1n(Y) + (1 - T)1n(1 - Y) \tag{5}$$

where, the number of the training data sets is expressed by N_D, T indicates the class outputs detected and Y displays the class outputs expected.

Methods used for validating the models

Receiver operating characteristics (ROC) curve. For any two distinct vectors where the first vector describes the binary presence-absence state of a particular action and the second vector gives the related probability predictions, the ROC curve can be prepared which is a common cut-off dependent diagnosis[42,43]. The overall model success can be recognized based on AUC values, as per the classification provided by Hosmer and Lemeshow[44]. There are four major elements of ROC curve: e true positive (P), false positive (Q), false negative (R) and true negative. The following measures including sensitivity or true positive rate (TPR), false positive rate (FPR), true negative rate or specificity (TNR), false negative rate or miss rate (FNR), efficiency, precision, negative predictive value (NPV), Matthews correlation coefficient (MCC) and Cohen's Kappa have been computed from these elements for validating the models:

$$\text{TPR} = \frac{P}{X} = \frac{P}{P + R} \tag{6}$$

$$\text{FPR} = \frac{Q}{Y} = \frac{Q}{Q+S} = 1 - \frac{S}{S+Q} \tag{7}$$

$$\text{TNR} = \frac{S}{Y} = \frac{S}{S+Q} \tag{8}$$

$$\text{FNR} = \frac{Q}{X} = \frac{Q}{Q+P} = 1 - TPR \tag{9}$$

$$\text{Efficiency}\left(\text{accuracy}\right) = \frac{P+S}{T} \tag{10}$$

where X, Y and T are the number of landslides, non-landslides and sum of total landslides and non-landslides, respectively

$$\text{Precision} = \frac{P}{P+Q} \tag{11}$$

$$\text{NPV} = \frac{S}{S+R} \tag{12}$$

$$\text{MCC} = \frac{(P \times S) - (Q \times R)}{\sqrt{(P+Q)(P+R)(S+Q)(S+R)}} \tag{13}$$

$$\text{Cohen's Kappa} = \frac{(P+S) - [(P+R)(P+Q) + (R+S)(Q+S)]/T}{T - [(P \times S))(P+Q) + (R+S)(Q+S]/T} \tag{14}$$

RMSE. The root mean square error (RMSE) was determined on the basis of the variations between the values expected by a model and the values actually observed (Eq. 9):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \left|\widetilde{P} - P\right|^2}{N}} \tag{15}$$

where N is sample size, $\widetilde{P}$ and P is predicted and observed values of dependent variable, respectively.

MAE. Mean absolute error (MAE) is measured as the sum of the differences between the expected value and the actual value of the model, without taking their direction into account:

$$MAE = \frac{1}{N} \sum_{i=1}^{n} \left|\widetilde{P} - P\right| \tag{16}$$

where N is sample size, $\widetilde{P}$ and P are predicted and observed values of dependent variable, respectively.

Relative landslide density (R-Index). R-index was used to evaluate the vulnerability maps of landslides. R-index is calculated as following[45].

$$R = ((xi/Xi)/\sum(xi/Xi)) \times 100 \tag{17}$$

where xi is the percentage of the area that is vulnerable to landslides in each vulnerability class, and Xi is the percentage of landslides in each vulnerability class. The maximum values of these vulnerability classes indicate the highest goodness-of-fit and excellent accuracy[46].

Results

Multicollinearity assessment. The findings of collinearity results indicate that no linearity exists among the LVCFs because the Tolerance and VIF values of these factors do not surpass their limits (Table 2) which suggests the aptness for inclusion in the modelling.

Results of IGR. The result of IGR method is shown in Fig. 8. The IGR values of the geo-environmental LVCFs for landslide vulnerability prediction were higher for geology, soil resistivity, rainfall, and elevation (Fig. 8a) and for socio-economic LVCFs, it is higher for distance to road and population density (Fig. 8b).

Landslide vulnerability analysis. Using the three models i.e. ANN, CNN, and DLNN and considering three aspects of vulnerability i.e. geo-environmental, socio-economic, and relative or overall vulnerability,

LVCFs	Collinearity statistics	
	Tolerance	VIF
Geo-environment factors		
Elevation	0.272	3.812
Slope	0.348	2.706
Aspect	0.898	1.113
Plan curvature	0.390	2.562
Profile curvature	0.282	3.547
Convergence index	0.517	1.933
Topographic position index	0.146	6.844
Terrain ruggedness index	0.270	4.286
Topographic wetness index	0.351	2.848
Valley depth	0.323	3.100
Relative slope position	0.327	3.059
Length of slope	0.181	5.521
Rainfall	0.269	3.723
NDVI	0.778	1.285
Distance from river	0.605	1.653
Socio-economic factors		
Population density	0.665	1.323
Old population density	0.273	3.646
House frequency	0.263	3.808
Literacy rate	0.233	4.286
Distance from road	0.266	3.765
Medical facility	0.627	1.596
Disability prevalence rate	0.123	8.114
Agriculture density	0.232	3.556
Number of the household for 30 min to reach road	0.334	2.269

Table 2. Collinearity results of LVCFs.

a total of 9 landslide vulnerability maps were derived which are ANN-LS (landslide susceptibility), ANN-SV (social vulnerability), ANN-RLV (relative landslide vulnerability), CNN-LS, CNN-SV, CNN-RLV, DLNN-LS, DLNN-SV, DLNN-RLV. Accepting natural breaks classification method, these landslide susceptibility maps were classified into five vulnerability zones: very low, low, moderate, high, and very high. The ANN-LS, CNN-LS, and DLNN-LS maps have 16.57%, 18.74%, and 17.75% area of the district as very high vulnerable for landslides from the scenario of geo-environmental setup. Similarly, ANN-SV, CNN-SV, and DLNN-SV maps have provided the socio-economic status of vulnerability as very high of 14.49%, 14.55%, and 14.40% area, respectively. In case of relative vulnerability maps ANN-RLV, CNN-RLV and DLNN-RLV have quantified 15.46%, 18.80%, and 14.66% of area as most vulnerable to landslides. The very high vulnerable zones have mostly occurred in the southern, south-west and south-east portion of the district and very low vulnerability is found in northern, north-east, and north-west regions. These results appear to be associated with the presence of weak geology, heavy rainfall during monsoon, and Phuentsholing-Thimphu national highway that passes through the first three portions of the district. Apart from this, rapid population growth in the above areas due to the fast development of Phuentsholing, the business city of Bhutan, resulted in new anthropogenic activities and thus weakening the soil (Fig. 9). The spatial extension of other vulnerability classes is given in Fig. 10.

Validation. Table 3 shows the comprehensive results of the ROC curve and other validation measures. The AUC computed using train data indicates success rate and AUC computed using test or validation data denotes prediction rate of the models[47]. Using the training data, AUC of ANN-LS, ANN-SV and ANN-RLV models are 0.893, 0.864, and 0.907 and using validation dataset these are 0.887, 0.846 and 0.902 respectively. The AUC using train and validation data are 0.912, 0.910 in CNN-LS; 0.893, 0.872 in CNN-SV and 0.921 and 0.928 in CNN-RLV models respectively. The AUC values of DLNN model are 0.880 (DLNN-LS), 0.867 (DLNN-SV) and 0.901 (DLNN-RLV) in the training case and 0.890 (DLNN-LS), 0.852 (DLNN-SV) and 0.900 (DLNN-RLV) in the validation case. Among these models, the CNN-RLV model has achieved the highest performance in terms of the AUC, accuracy, efficiency, kappa and predictive values and all other measures (Table 3). The result of reliability measures such as MAE and RMSE value is lowest for CNN-RLV followed by DLNN-RLV and ANN-RLV model. The R-index values for each model using geo-environmental aspect, socio-economic aspect or overall factors, is highest for very high vulnerability class followed by high vulnerability class (Table 4; Fig. 11).

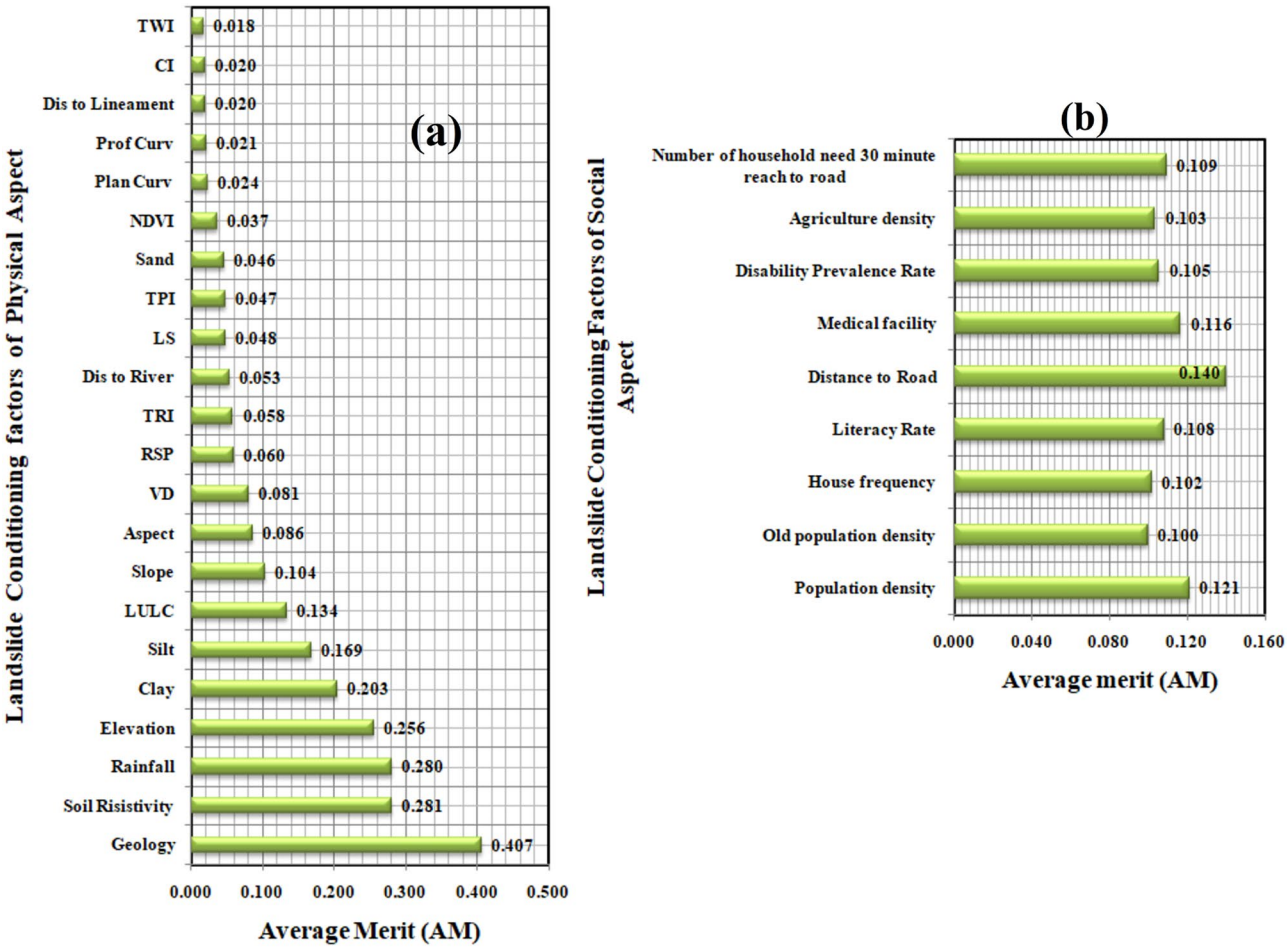


Figure 8. Average merit of LVCFs: (**a**) geo-environment factors, (**b**) socio-economic factors.

Factors importance analysis. The importance of the LVCFs in predicting landslide vulnerability was evaluated by random forest (RF) and chi-square attribute evolution technique (CSAE). The outcome of RF showed that the geology, rainfall, soil resistivity, and elevation were the most predictive factors for landslide vulnerability modelling in this research, followed by the sand and silt distribution and other geo-environment LVCFs (Fig. 12a). Among the socio-economic LVCFs, distance to road, population density factors have the most importance (Fig. 12c). The CSAE method yielded similar results to the RF model (Fig. 12b,d).

Discussions

Identifying or mapping the areas with future landslide potential is one of the most useful document for appropriate land use planning and mitigation decision making[13,47,48]. Evaluation of landslide susceptibility and vulnerability of a hilly area (as mountainous areas are subjected to the landslides) is indeed, necessary as it will serve as an essential dataset which will be used for identifying the points and places of relatively high landslide susceptibility and vulnerability which can be utilized for efficient and safe planning, design as well as the construction activities in an identified landslide zone. In the studied region, as well as other similar locations, there has been a significant growth in human socio-economic activities as well as other geo-environmental variables in recent decades. As a result, the frequency with which landslides occur in certain locations has risen. Such occurrences result in a significant loss of human life and property, as well as it results in a negative impact on different sectors such as tourism and infrastructure development[49]. A better working LVM has always been a subject of significant relevance within the community of numerous scholars across the world who are researching and investigating landslides. This is due to the fact that the methodology and conditioning factors used can have a substantial impact on the models' predictive performance[48]. It is very relevant to employ the mechanisms of factor selection to maximize the efficiency of landslide models by eliminating unwanted or trivial variables before training those[50]. CA and IGR method was adopted to accomplish this. In this study, CA revealed that all of the conditioning variables were effective and independent. In order to pick the most appropriate conditioning factor, there must be a non-collinear link between the LVCFs, which may be done via multicollinearity analysis[51]. Applying CA and IGR analysis, this study took into account a total of twenty-two geo-environmental and nine socio-economic factors, and the incorporation of these LVCFs has been justified in several studies[4,5,21]. It is very important to select LVCFs associated with past and present landslide events in vulnerability modelling. Although

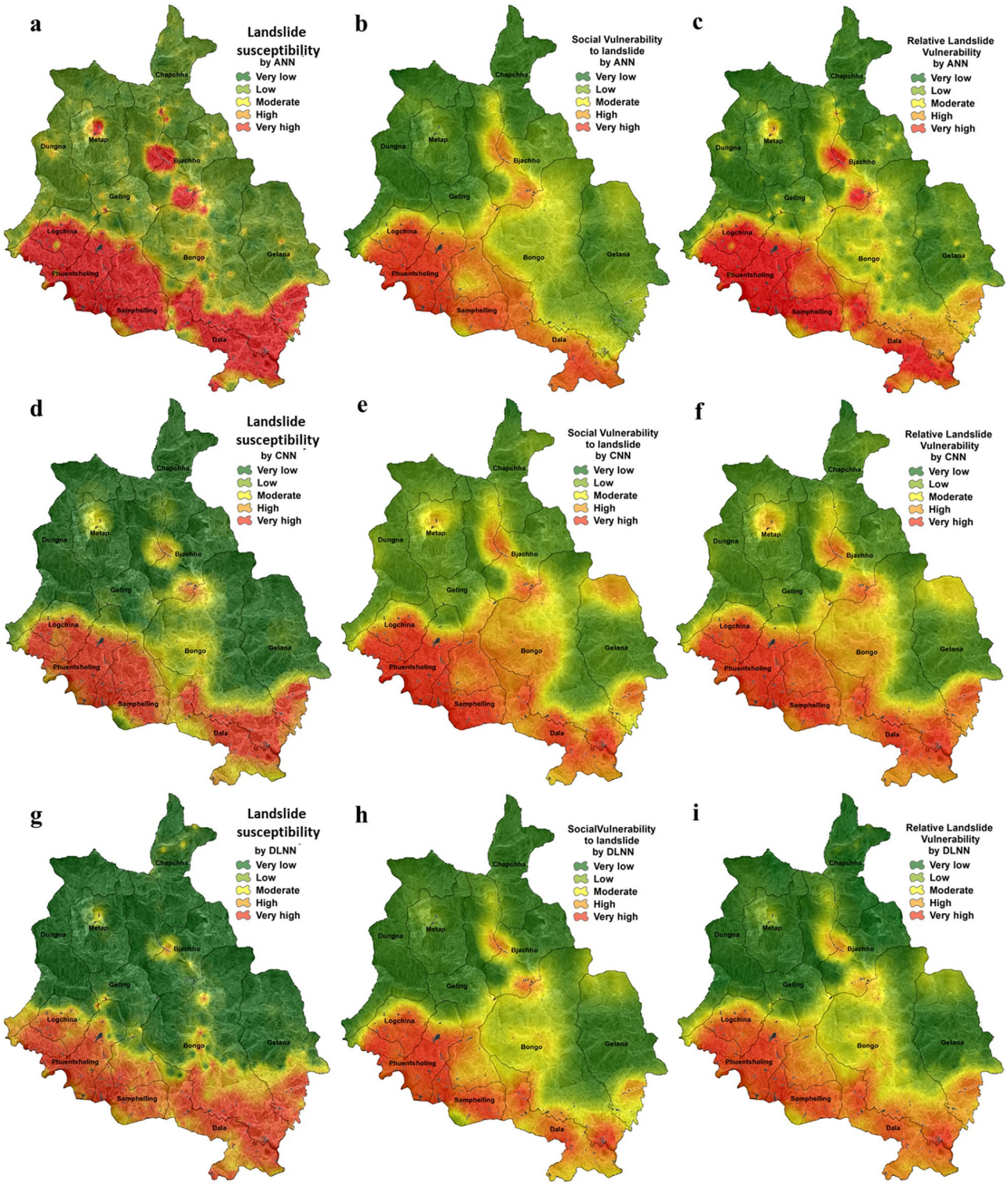


Figure 9. Physical, social and relative landslide vulnerability maps produced by: (**a**) ANN-LS, (**b**) ANN-SV, (**c**) ANN-RLV, (**d**) CNN-LS, (**e**) CNN-SV, (**f**) CNN-RLV, (**g**) DLNN-LS, (**h**) DLNN-SV and (**i**). DLNN-RLV.

landslides are quasi-natural hazard but when it results in destruction and harm to the economy and society, it indicates the association of socio-economic causes in the event happened. Therefore, to enhance the performance of modelling, this research was conducted taking into consideration both physical and social aspects; these factors are found to be very significant, especially in hillslope regions[52].

The study was approached towards mapping the landslide vulnerability considering geo-environmental and socio-economic aspects to improve the classification accuracy using deep learning and benchmark machine learning methods viz. CNN, DLNN and ANN. In all cases of vulnerability mapping (Physical, socio-economic

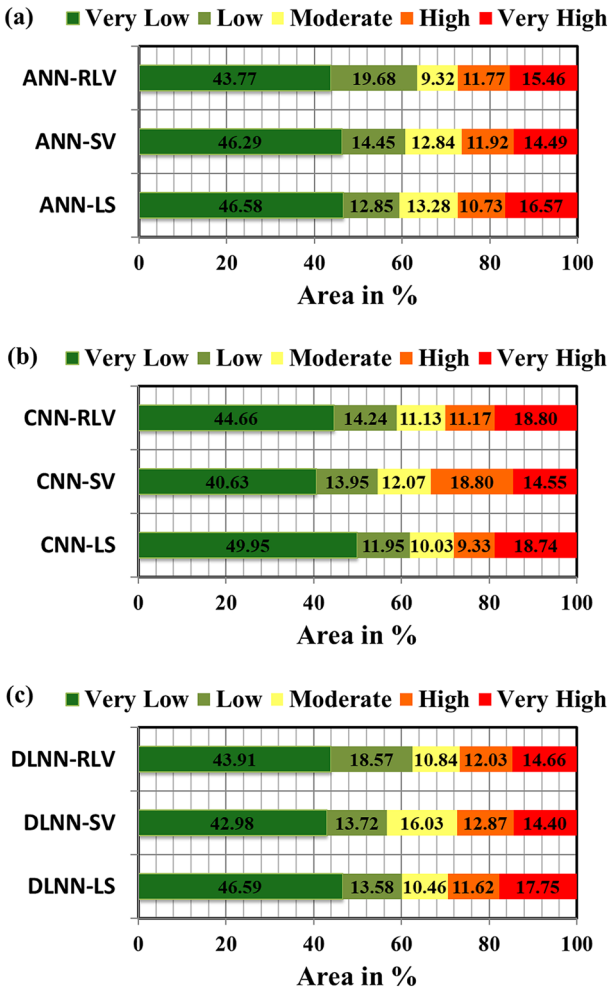


Figure 10. Distribution of physical, social and relative vulnerability classes produced using three models.

and combined) opted in this study, CNN has shown the best result followed by DLNN and ANN. In terms of the validation measures, the CNN-RLV model had the highest goodness-off-fit and excellent predictive performance, followed by the CNN-LS, CNN-SV, DLNN-LS, DLNN-S, DLNN- RLV, ANN-LS, ANN-SV and ANN-RLV models. Sadighi et al.[53] for landslide susceptibility assessment used MLP-NN with a Back-Propagation algorithm (BPANN), Adaptive Neuro-Fuzzy Inference System (ANFIS) models. However, result of models shows that the ANFIS-ICA had the superior results but ANN had quite good predictive accuracy i.e. AUC of 88.8%. In this study, ANN-RLV model has assured 88.7% success rate by the application of the AUC measure which is similar to the aforementioned study. In landslide susceptibility assessments, Bui et al.[9] used a DLNN model and compared its predictive efficiency with state-of-the-art machine learning models in Kon Tum province, Vietnam. Using ROC curve, the performance of the models revealed that the DLNN model had the highest goodness-of-fit and outperforming ANN, SVM model. Relatively better performance of DLNN than of ANN was also achieved in landslide vulnerability mapping for the present research. The efficiency of deep learning models compared to ML models was found to be better as per the study of Yao et al.[54] where authors have developed the deep neural network model based on semi-supervised analysis (SSL-DNN) for the landslide susceptibility estimation. For comparison, supervised models were introduced, including deep neural network (DNN), SVM, and logistic regression (LR). The result revealed that all comparable models were surpassed by the proposed SSL-DNN (AUC = 0.898) which is greatly supportive of outcome of the deliberating research. Application and enhanced competence of DLNN model can also be found in other hazard vulnerability and susceptibility mapping as Band et al.[55] proposed a DLNN model and an ensemble particle swarm optimization (PSO) algorithm with DLNN (PSO-DLNN), for gully erosion susceptibility mapping. These models were compared with ANN and SVM model. The PSO-DLNN model has the highest efficiency followed by the ANN and SVM. Therefore, the derived outcome of this research has a similar covenant to the very previous studies delegating relevance of adopted methodological scheme. Above all, r, the convolution neural network (CNN) model has provided superior results outperforming DLNN and ANN models in all of the social vulnerability, landslide susceptibility and relative or combined vulnerability assessment as exposed by all the adopted validation measures both in training set and testing set based analysis (Table 4). Therefore, accuracy of the DLNN model was better than the

Matrices	ANN-PV	ANN-SV	ANN-RLV	CNN-PV	CNN-SV	CNN-RLV	DLNN-PV	DLNN-SV	DLNN-RLV
Training dataset									
TPR	0.81	0.77	0.82	0.84	0.80	0.86	0.83	0.79	0.83
FPR	0.16	0.20	0.14	0.13	0.17	0.10	0.15	0.18	0.13
TNR	0.84	0.80	0.86	0.87	0.83	0.90	0.85	0.82	0.87
FNR	0.19	0.23	0.18	0.16	0.20	0.14	0.17	0.21	0.17
Accuracy	0.82	0.78	0.84	0.85	0.81	0.88	0.84	0.80	0.85
PPV	0.85	0.80	0.86	0.88	0.83	0.90	0.85	0.83	0.87
NPV	0.79	0.76	0.81	0.83	0.79	0.85	0.81	0.77	0.82
MCC	0.65	0.57	0.68	0.70	0.63	0.76	0.69	0.61	0.70
Cohen's Kappa	0.65	0.57	0.68	0.70	0.63	0.75	0.69	0.61	0.70
AUC	0.866	0.858	0.887	0.912	0.893	0.921	0.898	0.887	0.911
MAE	0.150	0.187	0.101	0.032	0.064	0.013	0.040	0.089	0.028
RMSE	0.387	0.432	0.318	0.180	0.254	0.114	0.199	0.299	0.169
Validation dataset									
TPR	0.82	0.82	0.86	**0.89**	**0.86**	0.92	0.88	0.85	0.90
FPR	0.11	0.17	0.10	0.09	0.12	0.07	0.09	0.15	0.09
TNR	0.89	0.83	0.90	0.91	0.88	0.93	0.91	0.85	0.91
FNR	0.18	0.18	0.14	0.11	0.14	0.08	0.12	0.15	0.10
Accuracy	0.85	0.83	0.88	0.90	0.87	0.92	0.90	0.85	0.90
PPV	0.90	0.83	0.90	0.92	0.89	0.93	0.92	0.85	0.90
NPV	0.81	0.82	0.85	0.89	0.86	0.92	0.88	0.85	0.91
MCC	0.71	0.65	0.75	0.81	0.74	0.85	0.79	0.71	0.81
Cohen's Kappa	0.71	0.65	0.75	0.81	0.74	0.85	0.79	0.71	0.81
AUC	0.857	0.846	0.879	0.910	0.872	0.928	0.890	0.852	0.900
MAE	0.069	0.103	0.044	0.023	0.027	0.011	0.024	0.067	0.022
RMSE	0.263	0.321	0.209	0.153	0.163	0.104	0.155	0.259	0.150

Table 3. Results of validation measures based on training and validation datasets.

conventional ANN machine learning technique. This is because greater number of samples and huge data could be handled by this model and the outcomes can be estimated with greater precision. The key benefit of deep learning is its formal system of self-governing DLNN layer organization learning. The conventional machine learning is incapable of processing such a vast number of inputs, and the result is less optimal in comparison to the deep learning model[55]. The accuracy of machine learning for different purposes was substantially enhanced in deep learning systems[55]. Yi et al.[50] developed a convolutional neural network (CNN) model for the spatial prediction of landslides. The result of CNN was compared with three conventional ML algorithms, i.e., logistic regression, multilayer perceptron (MLP) neural network and radial basis function (RBF) neural network which found CNN as the best fitted and excellent predictive model, followed by the MLP, logistic regression RBF. Wang et al.[39] applied the deep learning and ML models such as logistic regression, SVM, RF models for LS assessment. The result of that research also proved that CNN had the highest performance of predictive modelling followed by the ML models. With the agreement of the results of these studies, the present work also confirms the relatively highest adaptability of CNN model in deriving LVMs as reflected in the produced results of validation and accuracy measures. ReLU activation function was applied in the present study considering the aforesaid literature. The advantages of CNN model are that it considers all the neighbourhood information and can determine manifold stages of representations from input data[56.] It maintains the association of pixels using several factors and identifying internal elements[59].

Following Jenk's algorithm of natural breaks classification, LVMs were divided into five groups of susceptibility classes. This strategy of clustering data helps to reduce the mean–variance of each class from the mean within class range and to increase the discrepancy between each class from the means of the other classes[57]. From the analysis of the LVMs of this study, the very high vulnerability zone of landslide is found in the southern, south-western and south-eastern portion where the soil resistivity and geology is very weak. The amount of annual average precipitation is maximum in this part of the district. With very high elevation and steep slope, certain geological configurations influenced by socio-economic aspects tend to become unstable causing landslide. Phuentsholing is a highly urbanized centre located in the south upland slope with dense road network built through modification of the general slope that accelerates landslide processes.

Landslides are caused by a number of factors in a given area, but not all of them are equally responsible. During the field investigation, it was found that the landslides in the study region are caused by both natural causes (geological structure, heavy rainfall and very steep slope) as well as by human interferences (such as slope cutting for the construction of road, deforestation for expansion of agricultural land) which makes the area vulnerable to landslide. Relative landslide density index (R), in this work, helped to analyse the association

Models	Classes	No. of pixels	% of pixels	No. of landslides	% of landslides	R-index
ANN-P	Very low	5,607,364	46.58	21	6.11	2
	Low	1,547,452	12.85	45	12.86	15
	Moderate	1,598,676	13.28	47	13.50	16
	High	1,291,334	10.73	29	8.36	12
	Very high	1,994,494	16.57	207	59.16	55
ANN-S	Very low	5,573,335	46.29	38	10.93	3
	Low	1,739,451	14.45	21	6.11	6
	Moderate	1,546,377	12.84	82	23.47	27
	High	1,435,333	11.92	51	14.47	18
	Very high	1,744,824	14.49	158	45.02	46
ANN-RLV	Very low	5,269,576	43.77	1	0.32	0
	Low	2,369,894	19.68	6	1.61	1
	Moderate	1,121,903	9.32	33	9.32	14
	High	1,416,707	11.77	93	26.69	31
	Very high	1,861,241	15.46	217	62.06	54
CNN-P	Very low	6,013,570	49.95	1	0.32	0
	Low	1,438,915	11.95	8	2.25	3
	Moderate	1,207,156	10.03	20	5.79	9
	High	1,123,336	9.33	60	17.04	28
	Very High	2,256,343	18.74	261	74.60	61
CNN-S	Very low	4,891,668	40.63	2	0.64	0
	Low	1,678,914	13.95	7	1.93	2
	Moderate	1,453,602	12.07	32	9.00	11
	High	2,263,865	18.80	96	27.33	22
	Very high	1,751,272	14.55	214	61.09	64
CNN-RLV	Very low	5,376,679	44.66	1	0.32	0
	Low	1,714,734	14.24	5	1.29	1
	Moderate	1,339,692	11.13	20	5.79	8
	High	1,344,707	11.17	61	17.36	25
	Very high	2,263,507	18.80	263	75.24	65
DLNN-P	Very low	5,609,514	46.59	8	2.25	1
	Low	1,634,496	13.58	10	2.89	3
	Moderate	1,258,738	10.46	27	7.72	12
	High	1,399,513	11.62	55	15.76	21
	Very high	2,137,060	17.75	250	71.38	63
DLNN-S	Very low	5,173,934	42.98	5	1.29	0
	Low	1,651,690	13.72	8	2.25	2
	Moderate	1,929,658	16.03	32	9.00	8
	High	1,549,959	12.87	98	27.97	31
	Very high	1,734,078	14.40	208	59.49	59
DLNN-RLV	Very low	5,286,053	43.91	5	1.29	0
	Low	2,235,567	18.57	9	2.57	2
	Moderate	1,304,946	10.84	25	7.07	9
	High	1,448,229	12.03	102	29.26	33
	Very high	1,764,525	14.66	209	59.81	56

Table 4. Relative landslide density index of susceptibility classes of LSM models.

between produced vulnerability classes of landslide models and percentage of inventory landslides. The highest R-index value can be found in very high vulnerability class of each model followed by high vulnerability class which is positive for validating models.

The factor importance using RF (mean decrease Gini) and CSAE (average merit) represented the most important predictive LVCFs which are fittingly prominent in the southern part of the district (Fig. 12). Among the physical conditioning factors, both the RF and CSAE based evaluation has identified elevation, rainfall, geology, soil resistivity, soil clay and silt percentage etc. as the foremost persuasive factors for land sliding. The alike importance of the factors has reflected in several pieces of research such as elevation has been identified as a critical LVCF in various literature since most landslides occur in mountainous regions with a specific gradient[4,7,10]. Landslides are influenced by the structure, ordination, age, and exposure of the underlying surface[52]. The soil

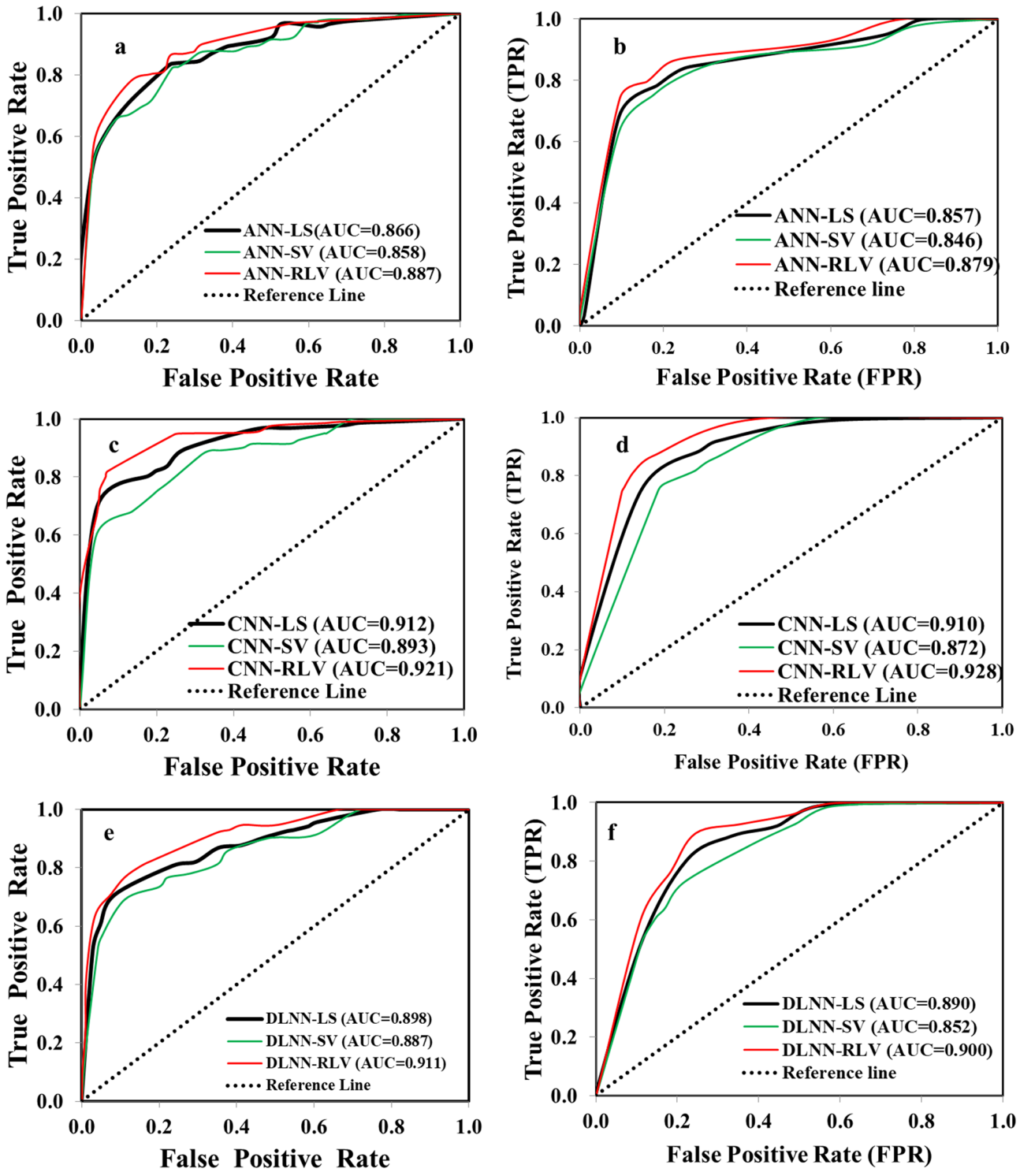


Figure 11. ROC curves used for validation of landslide vulnerability models: (**a**) ANN-LS, ANN-SV and ANN-RLV using Training, (**b**) ANN-LS, ANN-SV and ANN-RLV using validation datasets, (**c**) CNN-LS, CNN-SV and CNN-RLV using Training, (**d**) CNN-LS, CNN-SV and CNN-RLV using validation, (**e**) DLNN-LS, DLNN-SV and DLNN-RLV using training, (**f**) DLNN-LS, DLNN-SV and DLNN-RLV using validation datasets.

resistivity parameter has a positive connection with landslide susceptibility, showing that reducing soil resistance upturns the chance of a landslide, especially at higher elevations. Increased soil clay and silt content at medium to high altitude and upslope areas remain more unstable due to lower integrity than rocky sections, making it more vulnerable to seepage erosion, liquefaction, and fluidization. However, the type of soil and the amount of vegetation cover plays an important influence in this[46]. Among the socio-economic factors, RF and CSAE method have recognised distance to road, population density, agricultural density, house frequency factor as crucial

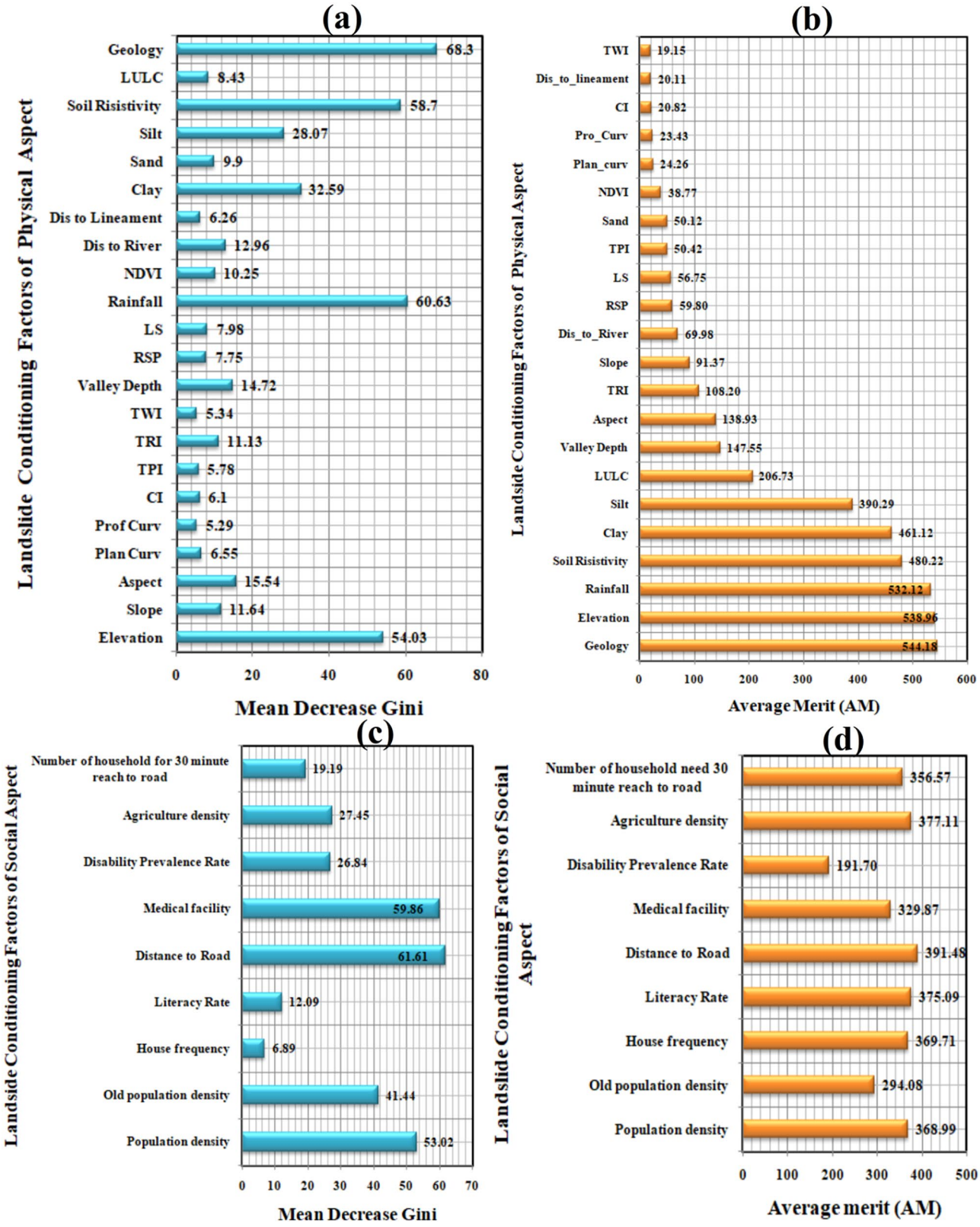


Figure 12. Importance of the LVCFs: (**a**) mean decrease Gini of geo-environmental LVCFs by RF, (**b**) average merit of geo-environmental LVCFs by CSAE, (**c**) mean decrease Gini of socio-economic LVCFs by RF, (**d**) average merit of socio-economic LVCFs by CSAE.

for making the area landslide vulnerable. A negative correlation exists for the distance to road LVF, indicating that the severity of possible landslide events increases in places as roads get closer, and vice versa. A substantial amount of work by Chan et al.[58], Weigand et al.[59] has established the role of these factors in triggering landslides.

Conclusion

Landslides have been seen in recent decades as the most critical natural risk that poses serious threat to both life and property all over the world. Thus short-term and long-term solutions are considered necessary to confront these daunting challenges. The landslide vulnerability map has recently become an important means of delineating landslide-prone regions and management. With the aid of sophisticated methods, proper data, and integration of remote sensing and a geographical information system, this can be accomplished. DLNN, ANN, and convolution neural network (CNN) models were used and r all aspects of a landslide event were considered which are novel approaches to perform the landslide vulnerability mapping of the district. Therefore, along with geo-environmental data, potential socio-economic factors were also considered as LVCFs using advanced factor selection techniques. CNN model achieved highest accuracy in modelling the vulnerability in the study area. As per the finding of the models, lower part of the district is highly susceptible and it needs immediate measures for managing. For the landslide risk supervision in the present and also for the future, the LVM can suggest implementing different management strategies like afforestation, barrier construction, and proper land use planning.

Received: 8 February 2021; Accepted: 3 August 2021
Published online: 12 August 2021

References

1. Thongley, T. & Vansarochana, C. Landslide susceptibility assessment using frequency ratio model at Ossey watershed area in Bhutan. *Eng. Appl. Sci. Res.* **48**(1), 56–64 (2021).
2. Kashyap, R., Pandey, A. C. & Parida, B. R. Spatio-temporal variability of monsoon precipitation and their effect on precipitation triggered landslides in relation to relief in Himalayas. *Spat. Inf. Res.* https://doi.org/10.1007/s41324-021-00392-8 (2021).
3. Nor Diana, M. I., Muhamad, N., Taha, M. R., Osman, A. & Alam, M. Social vulnerability assessment for landslide hazards in Malaysia: A systematic review study. *Land* **10**(3), 315 (2021).
4. Ram, P. & Gupta, V. Landslide hazard, vulnerability, and risk assessment (HVRA), Mussoorie township, Lesser Himalaya, India. *Environ. Dev. Sustain.* https://doi.org/10.1007/s10668-021-01449-2 (2021).
5. Kumar, P., Mital, A., Ray, P. C. & Chattoraj, S. L. Landslide hazard and risk assessment along nh-108 in parts of Lesser Himalaya, Uttarkashi, using weighted overlay method. In *Geohazards* (eds Gali, M. L. & Raghuveer-Rao, P.) 163–180 (Springer, 2021).
6. Li, Y., Chen, L., Yin, K., Zhang, Y. & Gui, L. Quantitative risk analysis of the hazard chain triggered by a landslide and the generated tsunami in the Three Gorges Reservoir area. *Landslides* **18**(2), 667–680 (2021).
7. Li, Z., Deng, X. & Zhang, Y. Evaluation and convergence analysis of socio-economic vulnerability to natural hazards of Belt and Road Initiative countries. *J. Clean. Prod.* **282**, 125406 (2021).
8. van Westen, C. J., Fonseca, F., & Van den Bout, B. Challenges in analyzing landslide risk dynamics for risk reduction planning. (2021).
9. Tsangaratos, P., Loupasakis, C., Nikolakopoulos, K., Angelitsa, V. & Ilia, I. Developing a landslide susceptibility map based on remote sensing, fuzzy logic and expert knowledge of the Island of Lefkada, Greece. *Environ. Earth Sci.* **77**, 363. https://doi.org/10.1007/s12665-018-7548-6 (2018).
10. Dikshit, A., Sarkar, R., Pradhan, B., Acharya, S. & Alamri, A. M. Spatial landslide risk assessment at Phuentsholing, Bhutan. *Geosciences* **10**(4), 131. https://doi.org/10.3390/geosciences10040131 (2020).
11. Reichenbach, P., Rossi, M., Malamud, B., Mihir, M. & Guzzetti, F. A review of statistically-based landslide susceptibility models. *Earth-Sci. Rev.* **180**, 60–91. https://doi.org/10.1016/j.earscirev.2018.03.001 (2018).
12. Bui, D. T., Tsangaratos, P., Nguyen, V. T., Van Liem, N. & Trinh, P. T. Comparing the prediction performance of a Deep Learning Neural Network model with conventional machine learning models in landslide susceptibility assessment. *CATENA* **188**, 104426. https://doi.org/10.1016/j.catena.2019.104426 (2020).
13. Pourghasemi, H. R. & Kerle, N. Random forests and evidential belief function-based landslide susceptibility assessment in Western Mazandaran Province, Iran. *Environ. Earth Sci.* **75**, 185. https://doi.org/10.1007/s12665-015-4950-1 (2016).
14. Chen, W., Sun, Z. & Han, J. Landslide susceptibility modeling using integrated ensemble weights of evidence with logistic regression and random forest models. *Appl. Sci.* **9**(1), 171. https://doi.org/10.3390/app9010171 (2019).
15. Pham, B. T. *et al.* Landslide susceptibility modeling using Reduced Error Pruning Trees and different ensemble techniques: Hybrid machine learning approaches. *CATENA* **175**, 203–218. https://doi.org/10.1016/j.catena.2018.12.018 (2019).
16. Liu, Y. & Wu, L. Geological disaster recognition on optical remote sensing images using deep learning. *Procedia Comput. Sci.* **91**, 566–575. https://doi.org/10.1016/j.procs.2016.07.144 (2016).
17. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Net.* **61**, 85–117. https://doi.org/10.1016/j.neunet.2014.09.003 (2015).
18. Yu, H., Ma, Y., Wang, L., Zhai, Y. & Wang, X. A landslide intelligent detection method based on CNN and rsg_r. In *Proceedings of the 2017 IEEE International Conference on Mechatronics and Automation (ICMA)*, Takamatsu, Japan, 6–9 August 2017, 40–44 (IEEE, 2017).
19. Ghorbanzadeh, O. *et al.* Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. *Remote Sens.* **11**, 196. https://doi.org/10.3390/rs11020196 (2019).
20. Pourghasemi, H. R. & Rahmati, O. Prediction of the landslide susceptibility: Which algorithm, which precision?. *CATENA* **162**, 177–192. https://doi.org/10.1016/j.catena.2017.11.022 (2018).
21. Sarkar, R. & Dorji, K. Determination of the probabilities of landslide events—A case study of Bhutan. *Hydrology* **6**, 52 (2019).
22. Gariano, S. L. *et al.* Automatic calculation of rainfall thresholds for landslide occurrence in Chukha Dzongkhag, Bhutan. *Bull. Eng. Geol. Environ.* **78**, 4325–4332 (2019).
23. Kuenza, K., Dorji, Y. & Wangda, D. Landslides in Bhutan. In *Proceedings of the SAARC Workshop on Landslide Risk Management in South Asia*, Thimphu, Bhutan, 11–12 May 2010, 73–80 (2010).
24. Yilmaz, C., Topal, T. & Suzen, M. L. GIS-based landslide susceptibility mapping using bivariate statistical analysis in Devrek (Zonguldak-Turkey). *Environ. Earth Sci.* **65**, 2161–2178. https://doi.org/10.1007/s12665-011-1196-4 (2012).
25. Cama, M., Lombardo, L., Conoscenti, C. & Rotigliano, E. Improving transferability strategies for debris flow susceptibility assessment: Application to the Saponara and Itala catchments (Messina, Italy). *Geomorphology* **288**, 52–65. https://doi.org/10.1016/j.geomorph.2017.03.025 (2017).

26. O'Brien, R. M. A caution regarding rules of thumb for variance inflation factors. *Qual. Quant.* **41**(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6 (2007).
27. Quinlan, J. R. *C4.5: Programs for Machine Learning* (Morgan Kaufmann, 1993).
28. Hosseinalizadeh, M. *et al.* Spatial modelling of gully headcuts using UAV data and four best-first decision classifier ensembles (BFTree, Bag-BFTree, RS-BFTree, and RF-BFTree). *Geomorphology* **329**, 184–193. https://doi.org/10.1016/j.geomorph.2019.01.006 (2019).
29. Pham, B. T., Bui, D. T., Prakash, I. & Dholakia, M. B. Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. *CATENA* **149**, 52–63. https://doi.org/10.1016/j.catena.2016.09.007 (2017).
30. Kavzoglu, T. & Mather, P. M. The use of backpropagating artificial neural networks in land cover classification. *Int. J. Remote Sens.* **24**(23), 4907–4938. https://doi.org/10.1080/0143116031000114851 (2003).
31. Haykin, S. *Neural Networks: A Comprehensive Foundation* (Prentice Hall, 1998).
32. Bergmeir, C. & Benitez, J. M. Neural networks in R using the stuttgart neural network simulator: RSNNS. *J. Stat. Softw.* https://doi.org/10.1863/jss.v046.i07 (2012).
33. LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **86**(11), 2278–2324 (1998).
34. Ding, A., Zhang, Q., Zhou, X. & Dai, B. Automatic recognition of landslide based on CNN and texture change detection. In: *2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)* 444–448 (IEEE, 2016).
35. Nair, V. & Hinton, G.E. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, 807–814 (2010).
36. Szegedy, C., et al. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1–9 (2015).
37. Zeiler, M. D. & Fergus, R. *Computer Vision–ECCV 2014: 13th European Conference*, Zurich, Switzerland, September 6–12, 2014. Proceedings, Part I, Vol. 818, p. 833 (2014).
38. Simonyan, K. & Zisserman, A., Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014).
39. Wang, Y., Fang, Z. & Hong, H. Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. *Sci. Total Environ.* **666**, 975–993. https://doi.org/10.1016/j.scitotenv.2019.02.263 (2019).
40. Kim, P. Matlab deep learning. *Mach. Learn. Neural Netw. Artif. Intell.* **130**, 21 (2017).
41. Lewis, N. D. C. Deep learning made easy with R: A gentle introduction for data science. In *Advances in Swarm Intelligence11th International Conference, ICSI (AusCov)*, Belgrade (2016).
42. Rahmati, O. *et al.* PMT: New analytical framework for automated evaluation of geo-environmental modeling approaches. *Sci. Total Environ.* **664**, 296–311. https://doi.org/10.1016/j.scitotenv.2019.02.017 (2019).
43. Lombardo, L. *et al.* Geostatistical modeling to capture seismic-shaking patterns from earthquake-induced landslides. *J. Geophys. Res. Earth Surf.* **124**, 1958–1980. https://doi.org/10.1029/2019JF005056 (2019).
44. Hosmer, D. W. & Lemeshow, S. *Applied Logistic Regression* (Wiley, 2000).
45. Baeza, C. & Corominas, J. Assessment of shallow landslide susceptibility by means of multivariate statistical techniques. *Earth Surf. Process. Landforms* **26**(12), 1251–1263. https://doi.org/10.1002/esp.263 (2001).
46. Meena, S. R., Ghorbanzadeh, O. & Blaschke, T. A comparative study of statistics-based landslide susceptibility models: A case study of the region affected by the gorkha earthquake in nepal. *ISPRS Int. J. Geo Inf.* **8**(2), 94. https://doi.org/10.3390/ijgi8020094 (2019).
47. Park, S., Choi, C., Kim, B. & Kim, J. Landslide susceptibility mapping using frequency ratio, analytic hierarchy process, logistic regression, and artificial neural network methods at the Inje area, Korea. *Environ. Earth Sci.* **68**(5), 1443–1464. https://doi.org/10.1007/s12665-012-1842-5 (2013).
48. Bui, D. T. *et al.* A novel hybrid evidential belief function-based fuzzy logic model in spatial prediction of rainfall-induced shallow landslides in the Lang Son city area (Vietnam). *Geomat. Nat. Haz. Risk* **6**(3), 243–271 (2015).
49. Chen, W., Pourghasemi, H. R., Kornejady, A. & Zhang, N. Landslide spatial modeling: Introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques. *Geoderma* **305**, 314–327 (2017).
50. Yi, Y., Zhang, Z., Zhang, W., Jia, H. & Zhang, J. Landslide susceptibility mapping using multiscale sampling strategy and convolutional neural network: A case study in Jiuzhaigou region. *CATENA* **195**, 104851. https://doi.org/10.1016/j.catena.2020.104851 (2020).
51. Lin, G. F., Chang, M. J., Huang, Y. C. & Ho, J. Y. Assessment of susceptibility to rainfall-induced landslides using improved self-organizing linear output map, support vector machine, and logistic regression. *Eng. Geol.* **224**, 62–74 (2017).
52. Roy, J., Saha, S., Arabameri, A., Blaschke, T. & Bui, D. T. A novel ensemble approach for landslide susceptibility mapping (LSM) in Darjeeling and Kalimpong districts, West Bengal, India. *Remote Sens.* **11**(23), 2866. https://doi.org/10.3390/rs11232866 (2019).
53. Sadighi, M., Motamedvaziri, B., Ahmadi, H. & Moeini, A. Assessing landslide susceptibility using machine learning models: A comparison between ANN, ANFIS, and ANFIS-ICA. *Environ. Earth Sci.* **79**(24), 1–14. https://doi.org/10.1007/s12665-020-09294-8 (2020).
54. Yao, J. *et al.* Assessment of landslide susceptibility combining deep learning with semi-supervised learning in Jiaohe County, Jilin Province, China. *Appl. Sci.* **10**(16), 5640. https://doi.org/10.3390/app10165640 (2020).
55. Band, S. S. *et al.* Novel ensemble approach of Deep Learning Neural Network (DLNN) model and Particle Swarm Optimization (PSO) algorithm for prediction of gully erosion susceptibility. *Sensors* **20**(19), 5609. https://doi.org/10.3390/s20195609 (2020).
56. Zhang, L., Xiao, T., He, J. & Chen, C. Erosion-based analysis of breaching of Baige landslide dams on the Jinsha River, China, in 2018. *Landslides* **16**(10), 1965–1979. https://doi.org/10.1007/s10346-019-01247-y (2019).
57. Jenks, G. F. The data model concept in statistical mapping. *International Yearbook of Cartography* **7**, 186–190 (1967).
58. Chan, H., Chang, C. C., Chen, P. & Lee, J. T. Using multinomial logistic regression for prediction of soil depth in an area of complex topography in Taiwan. *CATENA* **176**, 419–429 (2019).
59. Wiegand, C., Kringer, K., Geitner, C. & Rutzinger, M. Regolith structure analysis—a contribution to understanding the local occurrence of shallow landslides (Austrian Tyrol). *Geomorphology* **183**, 5–13 (2013).

Acknowledgements

The authors sincerely thank the various agencies of the Royal Government of Bhutan for providing rainfall data and the Border Roads Organisation, Government of India (DANTAK project) for providing landslide data. The authors acknowledge the two anonymous reviewers for their useful comments and suggestions.

Author contributions

S.S., R.S.: Methodology, format analysis, presentation, original draft preparation, review and editing; J.R., T.K.H.: Format analysis, software, original draft preparation; S.A., G.T.: Methodology, field investigation; D.D.: Field investigation, review and editing.

Funding
The present research was supported by International Science Council under Grant ISCROAP/IRDR/SG/2019/009.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to R.S.

Reprints and permissions information is available at www.nature.com/reprints.

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

Meta Analysis of Deep Learning Models for Doodle Recognition

Prateek Sethi
Department of Applied Mathematics
Delhi Technological University
Delhi, India
prateeksethi_2k17mc78@dtu.ac.in

Sachin Duhan
Department of Applied Mathematics
Delhi Technological University
Delhi, India
sachinduhan_2k17mc87@dtu.ac.in

Sanyam Gupta
Department of Applied Mathematics
Delhi Technological University
Delhi, India
sanyamgupta_2k17mc94@dtu.ac.in

Goonjan Jain
Department of Applied Mathematics
Delhi Technological University
Delhi, India
goonjanjain@dtu.ac.in

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498373

***Abstract* — Digital life has augmented the human horizons, code-driven systems have spread to more than half of the world's inhabitants owing to ambient information and connectivity. Successful combination & integration of human intelligence and machine, also known as Artificial Intelligence continues to impact the future of virtually every industry. From the Harappan civilization till the 21st century, visuals have played a crucial role in our life. The use of hand-made sketches is both frequent and persistent, which empowers human beings with greater retention and the free exchange of ideas and information. This mode of communication is both effective and helpful as witnessed across various communities. Considering the noisiness of the drawings, as the illustrators are often untrained and focus on ideas rather than the technicalities of art, doodle recognition is considered to have significant consequences in computer vision, Natural Language Processing, digit recognition, pattern recognition, and speech recognition. In this paper, we aim to implement various techniques of deep learning to analyze and classify doodles into predefined labels and estimate their efficacy to eventually improve classification accuracy. We have used data from the Google Quick Draw dataset to train and test the various Deep learning Models.**

Keywords — Convolution Neural Network; Deep Learning; Doodle Recognition; Transfer Learning; Image Recognition; AlexNet; DenseNet; GoogleNet; VGG19; MobileNet; SqueezeNet.

I. INTRODUCTION

Drawing and understanding pictures is a method of correspondence when words miss the mark or language is lacking due to various social and proficiency constraints. Individuals learning new dialects can likewise draw doodles and get interpretations. All these applications require computers to understand our doodles or quick line drawings.

Doodle recognition focuses on the intent with which the user is drawing, rather than modeling their unique, artistic styles. This foregoes the need to train the model on identifying distinctly styled sketches and vice-versa, the user may draw freely without any formal training. Identification and classification of these sketches still fall under the purview of image classification, which is a well-documented problem in the computer vision community. The grayscale nature of the sketches poses a further challenge when compared to colored-images classification. Further, the simplistic nature of drawings and the hugely diverse visual perception of objects by humans add to the dilemma. Another issue with drawing classification is the availability of large blank spaces throughout as compared to complete RGB photographs, which in a way can also help by allowing compression of the training data. Lastly, the sources of variation in both of these types of images-sketches and photographs are very different. Photographs show an impact of occlusion, viewpoint variation, illumination conditions, scale variation, which are of no consequence in hand-drawn doodles. The objective of the research work was inclined towards improving, and analyzing the accuracy of the model to predict the labels of the hand-made doodles, which has a wide range of applications. Ability to concede and classify the user made doodles has wide-ranging implications for the development of artificial intelligence in general. With the advent of new models and recent research in this domain, the improvement in performance of these new classifiers on the high noise dataset will surely impact the research directions in computer vision and pattern recognition. The remainder of this paper is organized as follows: Section II gives a detailed study on the various works done in the related field. Section III comprises the proposed framework of our model. In section IV the implementation and experiment details are discussed. Section V illustrates the evaluations performed and the results obtained.

II. RELATED WORK

The use of hand-made sketches is both frequent and persistent that empowers human beings with greater retention of ideas and information. This mode of communication is both effective and helpful as witnessed across various communities. Multiple approaches have been laid down by researchers for doodle & sketch recognition using CNN, RNN, HMM [1], CNN-RNN Hybrid & Transfer Learning. Neural Network Architectures for Sketches CNNs are a typical choice for feature extraction from sketches. Multiple Convolution Neural Networks methods have already been used for analyzing the static nature or the temporal sequential

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:53:36 UTC from IEEE Xplore. Restrictions apply.

property using RNNs. Sketch-a-Net [2] was the first CNN-based model with a sketch-specific architecture. It was directly inspired by AlexNet (Krizhevsky et al., 2012). Song et al. (2017) improved Sketch-a-Net further by adding spatial-semantic attention layers. A CNN-RNN hybrid architecture for sketches was proposed [3]. A novel Graph Neural Network (GNN) [4], the Multi-Graph Transformer (MGT) [5] was also designed for Free-Hand Sketch Recognition. Kim and Savarese examined SVM and KNN performance on image classification, specifically on airplanes, cars, faces, and motorbikes [6]. Lu and Tran architected a Convolutional Neural Network (CNN) to tackle sketch classification [7]. The recent introduction of numerous Deep Sketch models for sketch classification using CNN in 2017 by Seddati et al[8], was marked by the capability of the model to obtain an accuracy of 77.64% (Mean Average Precision-MAP) which was further improved to 79.18% as per TU-Berlin benchmark for sketches. They also added the residuals to further boost the accuracy to 93.02% on the sketchy dataset.[8]

A. Transfer Learning

In general, Transfer learning is an approach where the use available auxiliary data (that ecompasses the target training and test data) is used for training an existing model and used to solve the target problem. We have used the existing image recognition models that have already learnt patterns to solve a related but different problem, instead of from starting from scratch.[9] It was observed that this methodology required a lesser number of epochs, and hence less time taken for convergence, while increasing adaptability and flexibility of the model. Previous works include multiple ImageNet based Models used for the classification of flowers[10]. Such ImageNet models are now used in various image classification problems, which when used; provide better results when compared to the CNNs written from scratch. Transfer Learning is, therefore, an effective method for improving the performance of a model as the knowledge contained in the pre-trained model is applied to the model as can be seen in [11] and [12].

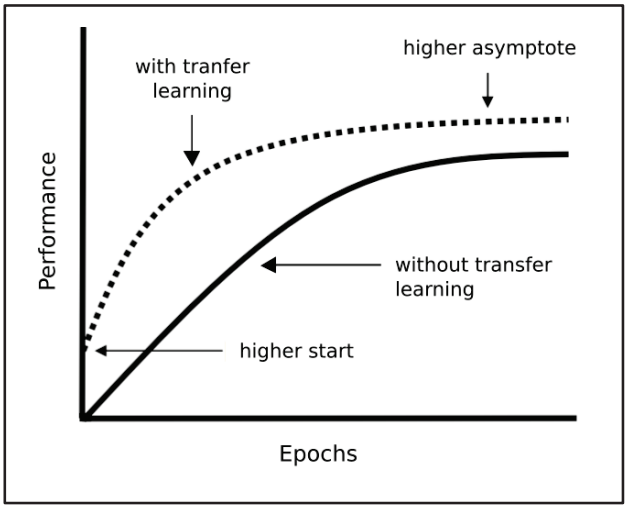


Fig. 1. Transfer Learning

III. EXPERIMENTAL SETUP

A. Dataset

Dataset is one of the most crucial sections of research that greatly impacts the overall performance of a model. There are numerous datasets available that can be useful during our study, dataset like WORDGUESS-160 (sketch-based word-guessing dataset for hand-drawn sketches), NIST Special Database 19 & MNIST dataset (binary images based dataset of handwritten digits recognition), sketchy database, SBIR dataset, Chars74K dataset, MSRA-TD500. ImageNet is another image database that can be used for training models, as it consists of 14 Million images that are organized in compliance to the WordNet hierarchy. For our research, we used a subset of google quickDraw dataset. The data has 25 categories of images, and the total number of images used was approximately 500000.

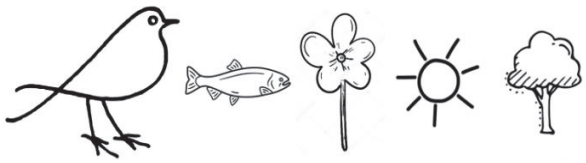


Fig. 2. Examples of the doodles available in the Quick Draw Dataset

This dataset complimented our requirements, as user-drawn doodles are a precise representation of human-to-human interactions and mode of exchange of information across various communities. The dataset proved to be highly worthwhile when it comes to comparing & differentiating the performance of various models like CNN, ResNet-18, DenseNet, VGG 19, AlexNet, MobileNet, and GoogleNet. We split the data in the ratio of 75:10:15 for the training, validation, and test sets for each label.

B. Data Preprocessing

Dataset consisted of multiple classes, where the doodles with labels were organized in different files. Firstly we shambled the PNG files, generating a compiled set of input-data from all classes which not just removed bias but also ensured randomness at the input layer for effective model training. The images in the data set were resized using an inbuilt MATLAB function. The images were transferred to the RGB scale by giving the grayscale value of the images to each R, B, and G parameter. This means that a pixel with the greyscale value of 150 is transferred to RGB channels as (150,150,150). This was a conscious choice since our dataset had only black and white images. In the later stage, we decided to pre-process the images by randomly rotating, flipping, altering and blocking some parts of it to add more noise, which in-return improved the ability of the model to handle noisy data.

C. Computational Resources

Training and validation play a crucial role in determining the results for any machine learning model. We used a 16 GB RAM with intel i7 10Gen processor, supported by GPU NVIDIA GeForce-GTX 3080 as hardware support for training the different models and approaches.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:53:36 UTC from IEEE Xplore. Restrictions apply.

IV. METHODOLOGY

In this study, we have tried to examine two approaches to image recognition: a simple Convolution Neural Network and Transfer Learning. For the simple Convolution Neural Network, we build a network from scratch and then train it on the dataset under study. For the Transfer Learning approach, we load a pre-trained Convolution Neural Network, i.e. a Convolution Neural Network already trained on a related, but different dataset. Then we replace the last fully connected and classification layers of the pre-trained network with layers in accordance to the dataset under study. The model is then trained on the dataset. The initial weights of the model, while generally chosen at random, are the weights corresponding to the dataset on which the network is previously trained, which results in better performance of the model and higher rate of convergence.

We applied transfer learning in the following two strategies: (i) We load a pre-trained network, and retrain it on the dataset under study. (ii) We freeze the first half of the total layers for each pre-trained model loaded, and train the remaining layers. In the former approach, training the pre-trained network on the dataset fine-tunes the parameters in accordance with the problem, while the network is trained as if from scratch, with the pre-trained weights as the initial weights The latter approach, on the other hand, improves the performance of the model by freezing the weights of initial layers and updating only the later layers, thereby reducing the number of operations to be carried out in each iteration. Most models extract only basic information in the initial layers, hence freezing the initial half of the layers allows for a balance of performance and accuracy. [13]. The same data set consisting of approximately 500000 images spanning across 25 classes was used for every model and each model was run for an equal number of epochs

A. Models

1) Simple CNN :

We Implemented a convolutional neural network for doodle classification, that was designed for handling the spatial inputs. CNN marked an important advance in the understanding of pictures and uses a three-dimensional framework where each set of neurons examines a particular area or image "feature". CNN architectures are difficult to design and develop from scratch as it involves selection of multiple hyperparameters and layer configurations. The end objective of moving to use a simple CNN model to ensure a lower performance threshold that highlights the various architectural aspects and to better compare the execution time of the model. Furthermore, to address the noise due to misleading training data, we added the larger batch size to the model architecture.

2) ResNet-18 :

It was observed that degradation increases as Deeper networks begin converging, a shallow network counterpart was used to replace the deeper network keeping other layers constant which reduced the error by a significant margin[14]. In other words, during back propagation, the information goes backwards, and the gradient inevitably passes through f(x), which may be convolution, batch normalization, or matrix multiplication. This phenomenon pertaining to back propagation results in trouble due to the various non-linearities involved and may make the gradient infinitely small. The ResNet 18 architecture makes use of an 'identity shortcut connection' which skips one or more layers. During the implementation of our first strategy we loaded a pre-trained version of ResNet-18, which is a deep residual learning framework that has learned rich feature classification for a wide range of images. The pre-trained ResNet network is represented as a Directed Acyclic Graph with 79 connections and 74 layers. While applying the second strategy, 37 layers were frozen and the rest of the layers were kept free for retraining.

3) GoogleNet :

GoogleNet uses a CNN inspired by LeNet but implements a novel element that is dubbed an inception module, which uses image distortions, batch normalization, and RMSprop. Some small convolution layers are used to minimize the number of parameters [15]. We observed that adding a wider and deeper inception network to the ensemble failed to improve the results considerably [16]. Due to empirical evidence suggesting that the impact of precise architectural parameters is comparatively minor, we omit the details of that network. The network was designed for computational efficiency and practicality, to eventually run inference on independent devices including even those with limited computational resources and low memory footprint. The pre-trained network is represented as a Directed Acyclic Graph with 144 layers and 170 connections out of which 72 layers were frozen for the second strategy.

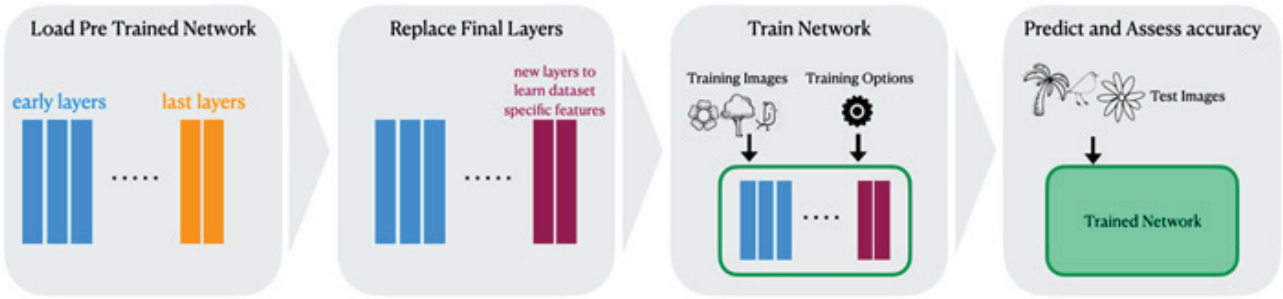


Fig. 3. Methodology for Transfer learning

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:53:36 UTC from IEEE Xplore. Restrictions apply.

4) AlexNet :

Developed by Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton [17], Alexnet is a deep CNN developed to classify the 1.2 million high-resolution images in the ImageNet dataset. Alexnet consists of 8 layers, with 5 convolutional and 3 fully connected layers. This model provided a drastic cut in the training time because it uses Rectified Linear Units (ReLU) instead of standard tanh function and Multiple GPUs that improve training time. This model also addressed the problem of overfitting through techniques like Data Augmentation and Dropout. For our 2 strategies, the first strategy used the pre-trained model in the form of a Series Network, and in the second strategy, 12 out of the total 25 layers were frozen to obtain the result.

5) MobileNet :

The model uses a technique of factorized convolutions which makes it light and quick. MobileNet facilitates the availability of the latest technologies like text, logo, and object recognition for users anytime and anywhere, irrespective of an Internet connection [18]. The use of depth-wise separable convolutions to ensure a lightweight deep neural network is the underlying feature of MobileNet architecture. This factorization, by breaking the interaction between the kernel size and number of output channels, not just improves the performance but also reduces the model size significantly. This model captures knowledge faster due to a smaller Neural network and fewer computations with the image matrix. In depth-separable convolution, we divide each input and filter channels first and pass the input to the respective filter, or in a depth-wise fashion. And afterward, piece together the output. We do point-wise convolution afterward, which is the same as the convolution of 1x1.

Adhering to our Transfer Learning strategies, we used the pre-trained MobileNet model, which is represented as a directed acyclic graph with 155 layers and 164 connections; on our pre-processed doodle dataset. We retrained the model in the first strategy, and froze 77 of the 155 layers while implementing the second strategy.

6) VGG 19 :

This is a variant of the Visual Geometry Group model, so named due to its design of 19 layers (16 convolution layers, 3 fully connected layers, 5 MaxPool layers, and 1 SoftMax layer). VGG has other variants including VGG11, and VGG16 among others. This network is distinguished by its simplicity, using in increasing depth only 3 omnipresent layers stacked on top of one another, and Max pooling functions to minimize capacity by volume. A softmax classifier is then followed by two fully connected layers.[19] By this we observe that deep CNNs can achieve considerably higher precision by implementing more layers, but the downside is that it requires a very large amount of data & is extremely expensive to train. To address this concern, we can replace a 7 * 7 convolution filter with layers of 3 * 3 convolution filters which reduces the number of parameters and impacts the computational time. We followed the above mentioned pre-processing methods that enabled us to feed in the target data to the input layers of the model. Kernels of (3*3) sizes with a 1 pixel stride size were also used in the model that enabled us to cover the whole conception of the doodles. we also used the spatial padding to preserve the spatial resolution of the image, followed by a max pooling over a 2*2 pixel window with stride 2. Furthermore, the introduction of the nonlinearity using the Rectified Linear Unit and use of sigmoid function to improve the computational time also resulted in better classification ability. The pre-trained VGG-19 network is represented as a Series Network with 47 layers and 23 of them were frozen while applying the second strategy.

7) DenseNet :

DenseNet architecture is a logical extension of ResNet. ResNet architecture fundamentally includes merging a previous layer into a future layer, where we force the network to learn residuals by the additive merges. The DenseNet concatenates outputs from the preceding layers rather than use addition. DenseNets have numerous advantages including strengthening feature propagation, and encouraging feature reuse apart from alleviating the vanishing gradient problem. They also substantially reduce the number of parameters.[20] Taking into account the depth of this network, the major concern was the ability of the model to propagate gradients backwards through all the layers in an effective manner. The performance of shallower networks on doodle recognition suggested that the features produced by layers in the middle impact the performance and classification accuracy. [16] This issue with denseNet was addressed by adding the auxiliary classifiers to the intermediate layers of the model. The pre-trained DenseNet network is represented as a Directed Acyclic Graph with 709 layers and 806 connections in total out of which 354 layers were frozen for our second strategy.

8) SqueezeNet :

SqueezeNet achieves an accuracy comparable to Alexnet on ImageNet with 50 times fewer parameters and a model size lesser than 0.5MB. This is achieved using three techniques which are- Make the network smaller by replacing 3x3 filters with 1x1 filters, Reduce the number of inputs for the remaining 3x3 filters and finally Downsample late in the network so that convolution layers have large activation maps.[21] With comparable accuracy which is being offered by the model, the smaller CNN architectures of the SqueezeNet offers the following advantages: (i) The model requires lesser bandwidth while updating from the cloud to an individual device. (ii) Smaller CNNs are viable to deploy on Field Programmable Gate Arrays and other hardware with limited memory (iii) Due to a Smaller Convolutional Neural Networks (CNNs), the computational resources and communication required during the training and testing is very less. The distinguishing factor for SqueezeNet is primarily due to the ability of the network to not learn the relationship between the classes, which is observed in models like VGG. On the contrary, in this architecture, an additional learning step is embedded within the transformations between various “fire modules”. The pre-trained network is represented as a Directed Acyclic Graph with 68 layers and 75 connections out of which 34 layers were frozen to obtain the result in case of the second strategy.

V. RESULT AND DISCUSSION

Table 1 summarises the results of our experiment with various models used during this research project, the result

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:53:36 UTC from IEEE Xplore. Restrictions apply.

pertaining to various models using multiple approaches is also shown in Fig 4. We observe that strategy 1 in which no layers are frozen when applied on GoogleNet gives the highest accuracy of 93.86% which is closely followed by ResNet 18 and AlexNet. This performance is considerably better than our simple CNN which has an accuracy of 87%. The results are in line with our expectations where the simple CNN has the least accuracy and other sophisticated models with more layers have better accuracy. While comparing the training time and accuracy of all the models under two strategies, some interesting observations are made. When half of the layers are frozen, the training time of GoogleNet improves by around 33% whereas the accuracy decreases by a mere 0.56%. This followed the similar trend for all models and freezing of layers resulted in reduction in accuracy.

TABLE I. DETAILS OF RESULT

Model	Accuracy
CNN	87%
AlexNet	Strategy 1 - 93.67%
	Strategy 2 - 88.08%
MobileNet	Strategy 1 - 90.12%
	Strategy 2 - 88.42%
VGG19	Strategy 1 - 92.78%
	Strategy 2 - 92.31%
Resnet18	Strategy 1 - 93.72%
	Strategy 2 - 90.97%
Dense Net	Strategy 1 - 93.36%
	Strategy 2 - 90.67%
GoogleNet	Strategy 1 - 93.86%
	Strategy 2 - 93.36%
Squeeze Net	Strategy 1 - 92.56%
	Strategy 2 - 91.28%

In the case of VGG19 and SqueezeNet, however, there is almost no effect in the training time of the model and the accuracy effectively remains the same with just a slight decrease. ResNet behaves differently where the accuracy decreases by 3% in the second strategy and the training time remains the same. AlexNet also shows a sharp decrease in accuracy of around 5% while the training time improves by around 5%. The DenseNet model shows a speedup of 25% with a decrease of 2.5% in its accuracy. We also observe that although they give higher accuracy, the training time for models like DenseNet, SqueezeNet, GoogleNet etc is also considerably higher than our simple CNN for both the strategies.



Fig. 4. Graphical representation of accuracy of different models

VI. CONCLUSION AND FUTURE WORK

In this research work, seven different deep learning models, using multiple strategies were analyzed based on their ability to accurately predict the label of the doodle dataset, a subset of Quick draw dataset. We employed the Transfer learning technique to obtain results, specifically using the architecture of the pre-trained networks with their respective pre-trained weights. A total of 15 different approaches were compared that can be used for doodle recognition; and among the various models discussed in this paper, the GoogleNet where all layers are retrained achieved the best results with an accuracy of 93.86%. It was observed that ResNet18, DenseNet, and GoogleNet performed approximately the same followed by VGG19 & SqueezeNet. Networks like GoogleNet, DenseNet have higher accuracies, but at the same time are computation-extensive. Lighter models like MobileNet and Squeezenet have shorter training times and require relatively less resources. Squeezenet is a network that requires a relatively low amount of resources and yet yields considerable accuracy, making it ideal for use in individual devices with low computation capabilities. The accuracy of the classification may further be improved by transferring the learned features of different models, in the future. The drawing stroke order of the doodle, if available, may also be incorporated to increase the efficacy of the models. Shallower and faster architectures, able to run on mobile devices appear to be the promising course of action for lightweight authentication mechanisms.

REFERENCES

[1] Tevfik Metin Sezgin, and Randall Davis, “HMM-Based Efficient Sketch Recognition”.

[2] S. J. Pan and Q. Yang, "A survey on transfer learning," in IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 10, pp. 1345-1359, Oct. 2010.

[3] Yu, Q., Yang, Y., Liu, F. et al. “Sketch-a-Net: A Deep Neural Network that Beats Humans”. Int J Comput Vis 122,411–425. 2019

[4] Jie Zhou,Gangu Cui,Zhengyan Zhang, “Graph Neural Networks: A Review of Methods and Applications". arXiv:1812.08434v4, 2019.

[5] Peng Xu 1 Chaitanya K. Joshi 1 Xavier Bresson 1, “Multi-Graph Transformer for Free-Hand Sketch Recognition”. arXiv:1912.11258v2, 2020.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:53:36 UTC from IEEE Xplore. Restrictions apply.

[6] Sarvadevabhatla, R. K., Kundu, J., and Babu R, V. “Enabling my robot to play Pictionary: Recurrent neural networks for sketch recognition”. 2016

[7] Kim, J., Kim, B. S., & Savarese, S. (2012). Comparing image classification methods: K-nearest-neighbor and support-vector machines. Ann Arbor, 1001, 48109-2122.

[8] Seddati, O., Dupont, S. & Mahmoudi, S. Multimed Tools Appl (2017) 76: 22333. https://doi.org/10.1007/s11042-017-4799-2

[9] Michael T. Rosenstein, Zvika Marx, Leslie Pack Kaelbling, “To Transfer or Not To Transfer”. 2005.

[10] Y. Wu, X. Qin, Y. Pan and C. Yuan, "Convolution neural network based transfer learning for classification of flowers," 2018 IEEE 3rd International Conference on Signal and Image Processing (ICSIP), Shenzhen, 2018, pp. 562-566

[11] M. Shaha and M. Pawar, "Transfer learning for image classification," 2018 Second International Conference on Electronics, Communication, and Aerospace Technology (ICECA), Coimbatore, 2018, pp. 656-660.

[12] L. Shao, F. Zhu, and X. Li, "Transfer learning for visual categorization: a survey," in IEEE Transactions on Neural Networks and Learning Systems, vol. 26, no. 5, pp. 1019-1034, May 2015

[13] Andrew Brock, Theodore Lim, & J.M. Ritchie, Nick Weston, “Freezeout: Accelerate Training by Progressively Freezing Layers”. arXiv:1706.04983v2, 2017

[14] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, “Deep Residual Learning for Image Recognition”. arXiv:1512.03385, 2015

[15] Vivek Paul1, Anshul Gupta, “Convolutional Neural Network in Medical Diagnosis”. 2019

[16] Christian Szegedy, Wei Liu, Yangqing Jia, “Going deeper with convolutions”. arXiv:1409.4842v1, 2014

[17] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton - ImageNet Classification with Deep Convolutional Neural Networks, Link - Here

[18] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam, “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications”. arXiv:1704.04861, 2017.

[19] Karen Simonyan, Andrew Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition”. arXiv:1409.1556, 2014

[20] Gao Huang, Zhuang Liu, Laurens van der Maaten, Kilian Q. Weinberger, “Densely Connected Convolutional Networks”. arXiv:1608.06993, 2016

[21] Forrest N. Iandola, Song Han, Matthew W. Moskewicz, Khalid Ashraf, William J. Dally, Kurt Keutzer, SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size”. arXiv:1602.07360, 2016

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:53:36 UTC from IEEE Xplore. Restrictions apply.

International Journal of Innovation Scientific Research and Review
Vol. 03, Issue, 07, pp.1475-1480, July, 2021
Available online at http://www.journalijisr.com

IJISRR

Research Article

ISSN: 2582-6131

METAHEURISTIC BASED TECHNIQUES FOR OPTIMAL SIZING AND SITING OF DISTRIBUTED GENERATIONS IN RADIAL DISTRIBUTION NETWORK

*** Prem Prakash**

Electrical Engineering Department, Delhi Technological University, Main Bawana Road, Shahabad Daulatpur, Delhi, India.

Received 13th May 2021; Accepted 14th June 2021; Published online 15th July 2021

ABSTRACT

An artificial intelligent (AI) based approach for optimal insertion and sizing of distributed generations (DGs) in distribution network is addressed in this research study. The objective function is constituted by considering different functions and this function is optimized by using AI based techniques. Further, basic objective of this study is to minimize the designed function under given constraints. The objective function is consisting system active power loss with DG and active power loss without DG, the ratio of bus voltage after DG to bus voltage before DG, and reciprocal of index of voltage stability. In this research article GA, PSO and GA-PSO hybrid based techniques are investigated to optimize the designed objective function. The proposed algorithms are applied on IEEE 69-bus test distribution network. From the simulation results it realized that all considered techniques produced almost similar results but computational time of GA is more further it is observed that GA-PSO based hybridized technique is better in terms of optimization of different parameters. It is concluded on basis of results that GA-PSO based technique is slightly better as compare GA, PSO in reference loss minimization, and system voltage improvement.

Keywords: Distributed Generations, Distribution network, Minimization of Network Loss, GA, PSO, GA-PSO based hybrid technique, DG.

INTRODUCTION

Commonly, all distribution networks are configured as radial structure. The configuration of such systems is very much simple as they are feeding by one source. In reference to distribution network the R/X (resistance to inductive reactance) ratio of these networks is more hence the power loss in such systems is more than that of transmission system. That is the reason the power loss of distribution systems needs to address in strategically in order to minimize. Further, in conventional radial distribution network the flow of power take place in unidirectional while the power flow becomes bidirectional as on when any other source like DGs are connected in the network as it is shown in Fig. 1. Therefore, it is crucial to minimize the power loss of distribution network when DGs are connected in these systems. Moreover, it is also equally important that the optimal size DG should be connected at optimal position otherwise it will lead towards increase the system losses and voltage instability.

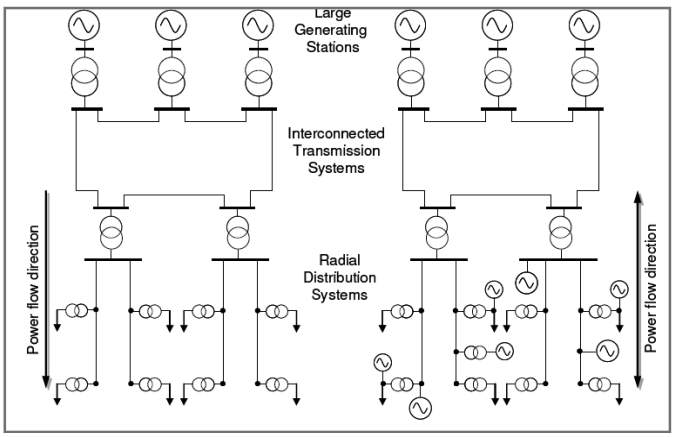


Fig.1: Structures of a typical electrical power system, (a) without DG (b) with DG

***Corresponding Author: Prem Prakash,**
Electrical Engineering Department, Delhi Technological University, Main Bawana Road, Shahabad Daulatpur, Delhi, India.

Since radial distribution networks are feeding by one source therefore the users which are connected at far end get reduced and fluctuating voltage profile due to load demand on the system is not remain constant and the requirement of reactive power is change continuously. The load profile is changing in nature which leads towards voltage instability at the most of the buses of the system [1] and [2]. Many of the researchers have found that distributed generators (DG) may be the possible solution of above mentioned drawbacks of distribution system if they are placed and sized optimally in distribution network [1], [2], [3], [4], [5], [6], [7], [8], [9], [12], [13], [14] [15], [16], [17]. DGs are competent to inject real and reactive component of power depending upon their types and these problems are arises due to excessiveness and scarcity of reactive power [3]. Further, DGs are competent to produce the power at low or without emission of greenhouse gases (GHG) which is mainly responsible for global warming. Further, DGs are emerging as an alternative source of power generation which has significant role in future in order to setting up new generation stations and expansion and restructuring of existing power generation units. According to available literature the power generation capacity of DGs will be 20% of total installed capacity across the world in future [4]. There are many reasons for increasing the infiltration level of DGs in distribution network which are given below

- Many DG technologies are existing in advanced and mature phase such as gas turbine, wind turbine generator (WTG), solar photovoltaic (SPV), fuel cell etc.
- The DG technologies are available in small range which is varies from few kW to few MW
- The installation of DGs technologies is easier one as it required less space and less investment
- The DGs units are placed at customer end therefore transmission and distribution costs are reduced or ignored.
- DG offers many advantages such as low environmental pollution, improved system bus voltage profile and reliability and minimize the system loss which in turn saving the cost of system.

- Generally, DGs are installed at close to load center therefore transmission and distribution costs are deferred

Related literature

This section of present research article highlights the work / investigations performed by several researchers for integration of DGs and renewable energy source (RES) in distribution networks. Pisică et al. [1] addressed a GA based technique for integration of DG in distribution network in optimal manner. The authors constituted an objective function in order to minimize the network power loss and investment cost. Moradi et al. [2] highlighted a differential evolution based optimization technique for optimal sizing and sitting of DG in distribution network for minimization of system power loss. Kansal et al. [3] investigated hybrid optimization technique for optimal sizing and sitting of different DGs to improve system voltage and minimize system loss. Salama [4] presented a review regarding various types of DG technologies and their definition. Zeinalzadeh et al. [5] addressed a novel particle swarm optimization (PSO) based method for optimal installation of DG for reduces network losses and improve bus voltage of network. Ghaffarzadeh and Sadeghi [6] highlighted a biogeography-based optimization approach for optimum placement of DG unit in distribution network in order to decrease system real power loss and enhance bus voltage. Authors of this research study also investigated the performance of PSO and GA approaches for considered function with reference to DG placement and sizing in distribution network. Gomez-Gonzalez et al. [7] proposed a novel SFLA based optimization method for probabilistic load flow for optimal installation of SPV type DG by considering uncertainties of solar radiation and obtained results are compared with Monte Carlo simulation based method. Devi and Geethanjali [10] applied a new and novel BFOA for installation of DG to diminish the system power loss and enhance bus voltage [10]. Khalesi et al. [11] have addressed the dynamic programming based approach for optimum installation and sizing of DG to decrease network power loss and improve system reliability by considering time varying load. Dulău et al. [12] highlighted the existence of major types of DG technologies and their optimization techniques for optimum sizing and placement in distribution network so that one can extract their potential in efficient manner. Roy and Sultana [13] projected a new QOTLBO approach for optimum allocation of DG to reduce network loss, improve index of voltage stability along with deviation of voltage in distribution network. Singh and Parida [14] offered a new method to categorize (rank wise) the node for installation of DG unit in order to diminish system power loss and enhance network bus voltage. Ramadan [16] proposed PSO based optimization technique for optimum integration of wind operated DG for minimization of system power loss and better bus voltage. Kumar et al. [17] applied a mixed integer non-linear programming (MINLP) based approach for optimal integration multiple DG units the objective of the research was to decrease the overall cost of fuel [17].

Formulation of problem

The reason of methodology presented in this research study is to optimize considered function which is combination of three functions. These functions include network losses, bus voltage and stability of voltage. Furthermore, prime and foremost objective is to combine all these functions to make a solitary function this function is known as fitness function and finally optimize this function [1], [2].

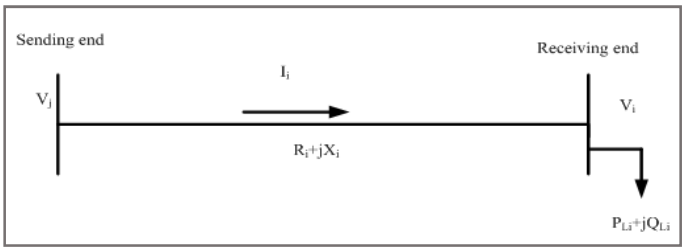


Fig. 2 Representation of two bus distribution network

Fig. 2 represents a two bus distribution network having two buses j and i. The X_i and R_i are the branch inductive reactance and resistance between bus j and i respectively

Objective function

In present research article the objective function is designed by taking three different functions as power loss of system, voltage profile of different buses of considered network and finally system stability is considered as third function. The description of the objective function is given as follows

$OF = Min\left(k_1 f_1 + k_2 f_2 + k_3 f_3\right)$ Where k_1, k_2 and k_3 are constants and f_1, f_2 and f_3 are defined as follows and $k_1+k_2+k_3 = 1$ is the weighting factor whose sum is equal to one

Network Power Loss

The active power loss of distribution network is a crucial and fundamental factor and need elaborate in detail and its description is explained as below. As it is factor of great importance therefore the priority is given as highest i.e. (0.5) to this factor. Here, the first function f_1 comprises network real power loss with DG and without DG. Therefore, f_1 may be represent as

$f_1 = \dfrac{\text{System power loss with DG}\left(P_{LDG}\right)}{\text{System power loss without DG}\left(P_L\right)}$ Where P_{LDG} and P_L are explained as under

$$P_{LDG} = \sum_{i=1}^{n-1}\left(I_i - DI_{dg}\right)^2 R_i$$

$P_L = \sum_{i=1}^{n-1}\left(I_i^2 R_i\right)$ Where

I_i = i^{th} branch current (in Ampere)
R_i = i^{th} branch resistance (in Ω)
I_{dg} = Injected current from DG (in Ampere)

Bus voltage profile

As on when DG unit is placed in distribution system it injects the current in distribution system. Due to this the magnitude of current get reduced which in turn reduces the system losses and improve system bus voltage. Hence, the voltage profile at each bus increased. i.e. ± 5% deviation of bus voltage is allowed. The, second function is f_2 and it is acknowledged as voltage deviation index (VDI). In projected scheme the goal is to minimize the VDF

Second function f_2 is

$f_2 = \dfrac{\sum_{i=1}^{n}\left(V_{bdi} - 1\right)}{\sum_{i=1}^{n}\left(V_i - 1\right)}$ Where V_{bd} is voltage magnitude when DG is

placed and V_i bus voltage before DG placement here n represents the buses number of considered network and 'i' varies from 1 to n

Voltage stability index (VSI)

If changes in load are significant the network is capable to retain voltage in tolerable limit without losing the synchronism is termed as voltage stability. As distribution systems are operated in voltage stability limits, if reactive power is not supplied in adequate amount the voltage stability limit of system is not remain in stable domain. Therefore, it is necessary to incorporate the VSI factor in objective function. The third function that is f_3 is represented in the form of voltage stability index (VSI), it is given as Further, I_i (branch current between bus j and k) is determined by applying following formula

$$I_i = \frac{V_j - V_k}{R_i + jX_i} \quad (1)$$

$$P_i + jQ_i = V_i I_i^* \text{ or } I_i = \frac{P_i - jQ_i}{V_i^*} \quad (2)$$

VSI is given as

$$VSI_i = |V_j|^4 - 4\left[p_i(i)R_i + q_i(i)X_i\right]|V_j|^2 - 4\left[p_i(i)R_i + q_i(i)X_i\right]^2 \quad (3)$$

For improvement of *VSI*, f_3 is given by reciprocal of *VSI*, therefore

$$f_3 = mean\left\{\sum_{i=2}^{n}\frac{1}{VSI_i}\right\} \quad (4)$$

For stability the value of $VSI_i > 0$, the value of *i* varies from *2, 3….n* and *n* indicates the total number of bus of the test distribution system.

Constraints

For load balancing constraints following equation must be satisfied

$$P_{IG} - P_{LI} - V_i\sum_{k=1}^{n}V_k Y_{ik}\cos\left(\partial_i - \partial_{ik} - \theta_{ik}\right) = 0 \quad (5)$$

$$Q_{GI} - Q_{LI} - V_i\sum_{k=1}^{n}V_k Y_{ik}\cos\left(\partial_i - \partial_{ik} - \theta_{ik}\right) = 0 \quad (6)$$

Where, i is total number of branches and varies from 1 to n-1 and n is total number bus in considered network

System voltage constraints, the node voltage is varies between lower and upper limit i.e. five percent variation is allowed in either side further node voltage not goes below minimum voltage and simultaneously not exceeds the maximum value of voltage

$V_i^{\min} < V_i < V_i^{\max}$ where, i = 1, 2 . . . . . n

Current constraints is given as

$I_i \le I_i^{\max}$; $i = 1, 2, 3.....n-1$ Here n is total number of bus

Moreover, three different functions which are described earlier are combined together in order to constitute the final objective function which is to optimized is expressed as follows

$$OF = \left[f_1 + f_2 + \frac{1}{f_3}\right] \quad (7)$$

The considered is optimized by applying by GA and PSO based optimization techniques

Optimal placement and sizing of DG using GA and PSO

The present research article is focused on optimal placement and optimal sizing of DGs in radial distribution network which are capable to inject active power is performed. In present paper the considered objective function is optimized by GA, PSO and GA-PSO hybridized based optimization technique.

Genetic Algorithm (GA)

GA starts searching from the set of population. It is a population based optimization technique in which random population is selected in order to find out random solution. This randomly generate group of solutions is known as chromosomes of the problem. The process of GA is described as follows

Step 1: First of all generate initial population in order to obtain a solution of n chromosomes and set time counter

Step 2: Determine fitness value of objective function for each solution in the initial population and search optimum value of considered function

Step 3: Updates the time counter

Step 4: Apply mutation and crossover and generate a new population which is assigned by user

Step 5: Now run algorithm for newly generated population

Step 6: If stopping criterion is met then stop the algorithm.

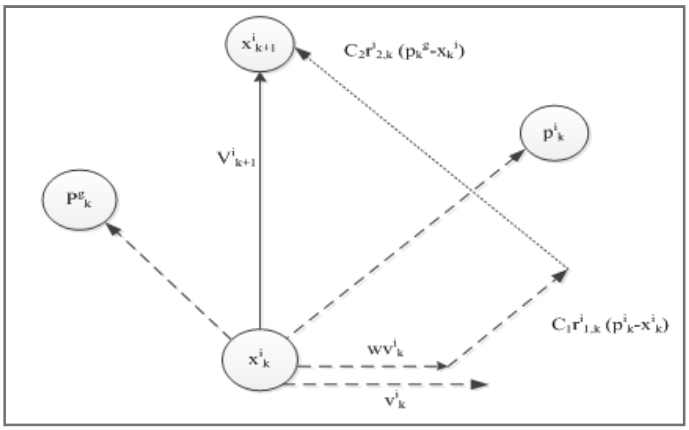


Fig.3 Position and velocity updating criterion for PSO based optimization technique

Fig.3 shows the updating and variation of searching point by PSO based technique. Further, for projected technique the maximum and minimum size of DG units considered in between 2×10^6 kW to 100 kW respectively.

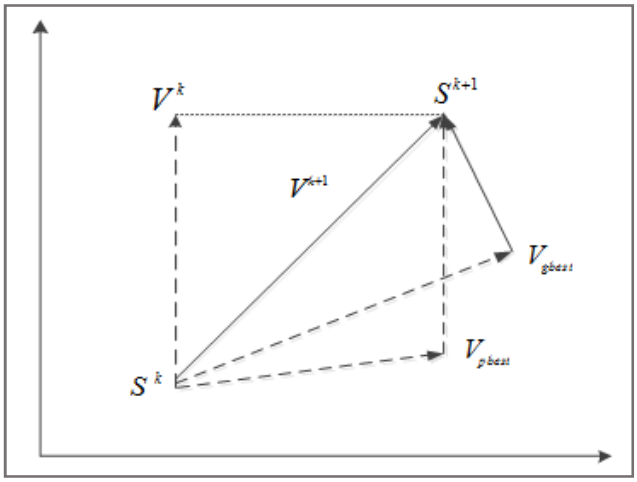


Fig. 4 Variation of searching points on the basis of PSO technique

Variation of points to be search in search-space to find out optimal location for DG placement in case of PSO based optimization technique is shown in Fig. 4

Description of PSO based algorithm

This is a nature inspired based optimization approach. The process of this approach is based on fish schooling and flocking of birds.
Following steps are involved in order to incorporate the PSO technique
Step 1: Initialize the population in order to find out appropriate solution
Step 2: Seek best values for objective function as P_{best}
Step 3: Renew time counter t = t +1
Step 4: Updates inertia weights
Step 5: Update particle velocities according to following equations

$$v_{i,it} = w \times v_{i,it-1} + c_1 \times rand_1 \times \left(pbest_{i,it-1} - x_{i,it-1}\right) + c_2 \times rand_2 \times \left(gbest_{i,it-1} - x_{i,it-1}\right) \times \left[1 + \frac{-range}{\max ite} \times (Ite - 1)\right]$$

$$x_{i,it} = x_{i,it-1} + v_{i,it}$$

Proposed methodology

In this proposed approach an appropriate installation and sizing of DG in order to optimize the fitness function which is made up by adding three different functions. The first function is constituted by taking real power loss before and after DG. Moreover, second function is also the ratio system bus voltage with DG and bus voltage without DG. Finally, third function is reciprocal of voltage stability index (VSI). The present research study is separated in two fold one is the optimum placement of DG unit and second part is the optimum sizing of DG. Moreover, in present research paper the minimum and maximum value of size of DG is 10 kW and 2 MW is considered to simulate the algorithm.

RESULT AND DISCUSSIONS

In the present paper the proposed algorithm is tested on IEEE 69-bus and 85-bus test distribution systems. Except root bus all buses are considered for DG placement. Further, in considered network which are consider as 69-bus, and 85-bus having bigger search space therefore the applied methods require more computational time which is major drawback of the technique.

IEEE 69-Bus Distribution System

Fig. 5 is used for representation of single- line diagram of IEEE 69-bus test radial distribution network. This system has a total peak load of (3.80219 + j2.69460) MVA. Table 1 represents the details of various results obtained by GA, PSO and hybrid optimization technique before and after DG placement for IEEE 69-bus distribution system. The GA based method provides bus number 62, 60 and 59 are obtained as candidate buses where DG could be placement. Moreover, their corresponding DGs sizes are 0.13267 MW, 1.52783 MW and 0.94396 MW. The method reduces system power loss from 225.07161 kW to 22.214 kW. Moreover, the value of different functions considered in proposed technique is obtained by GA as f_1 = 0.09870, f_2 = 0.34071 and f_3 = 1.00436 respectively. And optimal positions of buses where DG unit can be installed are obtained as 62, 60 and 59 respectively. Further, the various results obtained by PSO as bus number 24, 21 and 25 are obtained as candidate buses for DG placement. Their corresponding DGs sizes are 0.59092 MW, 1.00354 MW and 0.06320 MW respectively. The method reduces system power loss from 225.07161 kW to 20.304 kW. Moreover, the value of functions considered in proposed technique is obtained by PSO as f_1 = 0.09021, f_2 = 0.54175 and f_3 = 1.11099. Again, GA-PSO hybrid optimization method also produced similar kind of results as bus number 25, 15 and 27 are obtained as candidate buses for DG placement. Their corresponding DGs sizes are 1.18889 MW, 0.45816 MW and 1.08090 MW. The method reduces system power loss from 225.07161 kW to 18.180 kW. Moreover, the value of functions considered in proposed technique is obtained by GA-PSO hybrid as f_1 = 0.08077, f_2 = 0.57047 and f_3 = 1.06753. The detailed results are presented in Table.1 which are obtained by GA, PSO and GA-PSO hybrid based optimization techniques for IEEE 69-bus radial system. From these results it is concluded the power loss reduction capability of PSO based method is more as compare to other technique. Fig. 6 represent the convergence of objective function by GA based optimization technique and Fig. 7 represents the convergence of objective function by PSO based optimization technique. From these figures it is observed that PSO based optimization technique has faster rate of convergence as compare to GA based optimization technique which demonstrate that GA require more simulation time than PSO. Fig. 8 shows the system voltage profile before and after DG placement using GA, PSO and GA-PSO based optimization techniques. From this figure it is observed that GA-PSO based hybrid method improves system voltage profile more as compare to GA and PSO based technique. Fig. 9 represents the voltage stability index (VSI) at different bus which is evaluated by equation 3, from this figure it is observed that bus 1 has the highest value of VSI with 1.0 while bus 27 with 0.9840 has the lowest value of VSI in GA based optimization. Moreover, from this figure it is observed that bus 29 has the highest value of VSI with 1.01720 while bus 65 with 0.69608 has the lowest value by PSO based optimization. Further from Fig. 9 it is observed that minimum value of VSI is exists at bus number 27 with 0.8058 by GA-PSO based hybridized method. Moreover, it is observed that GA take more time for convergence as it converge after 100 iterations while PSO take lesser time for convergence as it converge after 20 iterations only. Additionally, system loss reduction capability of all three methods is almost similar but GA-PSO hybrid has slightly more loss reduction capability and voltage improvement capability. Again from this table it is noticed that the system loss reduced significantly by all these methods.

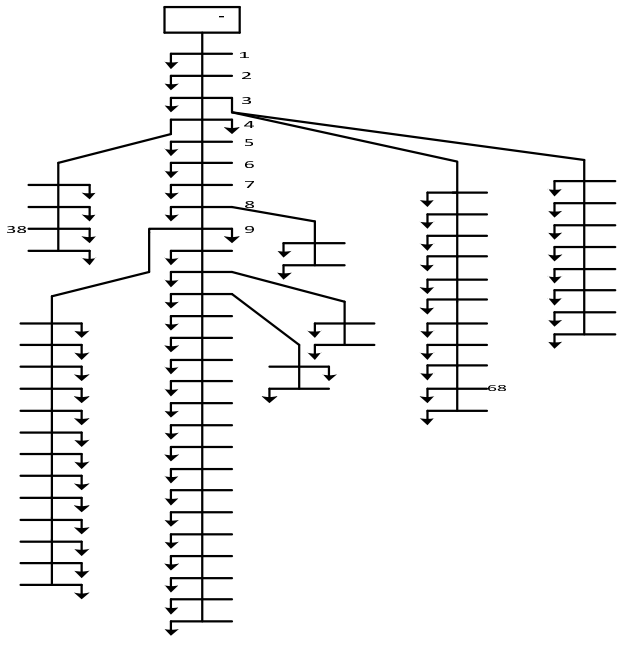


Fig. 5 Single line diagram for IEEE 69-bus radial distribution network

Table 1: Various results obtained by GA, PSO and GA-PSO based hybrid optimization technique for multiple DG placements for IEEE 69- bus test system

S.No.	Optimization technique	Candidate bus for DG placement	DG size (in MW)	System power loss before DG (in kW)	System power loss after DG (in kW)	Fitness function value
		62	0.13267			f_1=0.09870
1.	GA	60	1.52783	225.07161	22.21472	f_2=0.34070
		59	0.94396			f_3=1.00436
	PSO	24	0.59092			f_1=0.09021
2.		21	1.00354	225.07161	20.30404	f_2=0.54175
		25	0.06320			f_3=1.11099
		25	1.18889			f_1=0.08077
3.	GA-PSO Hybrid	15	0.45816	225.07161	18.18008	f_2=0.57047
		27	1.08090			f_3=1.06753

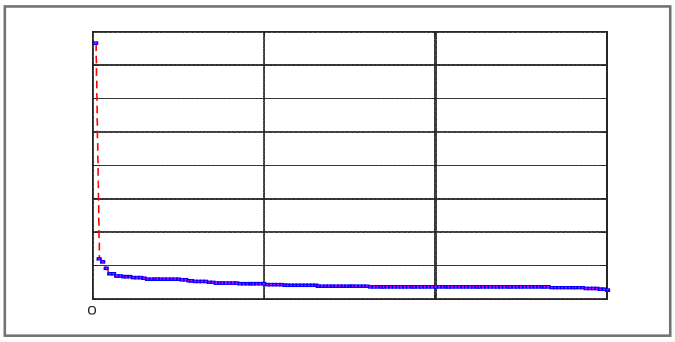


Fig.6 Variation of objective function by GA based optimization technique for IEEE- 69 bus

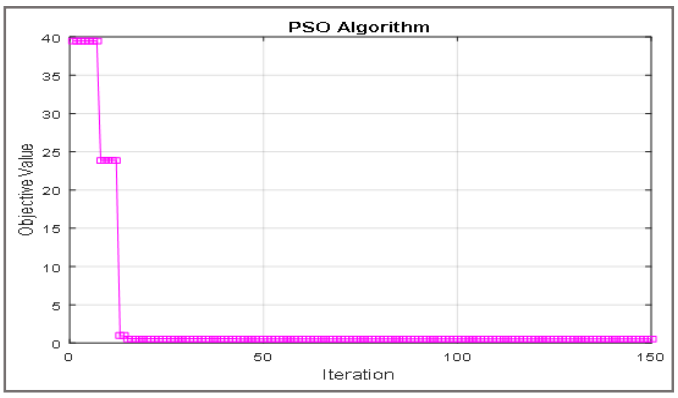


Fig.7 Convergence of objective function by PSO based optimization technique for IEEE 69-bus system

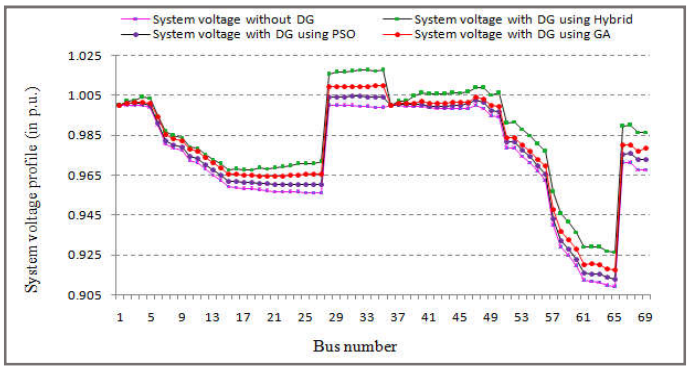


Fig.8 System voltage profile for IEEE-69 bus distribution network at different buses before and after DG placement by GA, PSO and hybrid based optimization technique

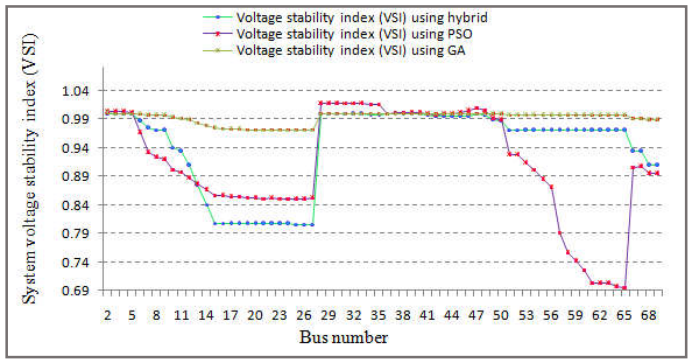


Fig.9. System voltage stability index (VSI) at various buses for IEEE 69-bus system by GA, PSO and GA-PSO based hybridized optimization technique

CONCLUSIONS

The present article focuses on optimum placement and optimum capacity of DG in distribution system is proposed. The objective function is designed by considering three functions. These function are as system loss, system voltage and system voltage stability index. The optimization of considered objective function by three different optimization techniques by GA, PSO, and GA-PSO hybrid optimization methods is the main goal of this research paper. On the basis simulation results it is observed that system loss reduced considerably and system voltage profile improved significantly. Additionally, GA, PSO and GA-PSO based hybridized method have been applied for optimization of DG locations and sizes. It is concluded that by means of achieved results that PSO, GA, and hybrid GA-PSO provide almost similar outcomes in reference of bus voltage enhancement and system loss minimization. But hybrid GA-PSO based optimization method is slightly additional efficient in respect of bus voltage improvement and system power loss reduction is concerned. Finally, it is concluded that GA-PSO based hybrid method provides better results.

REFERENCES

1. Pisic, C. Bulac, and M. Eremia, “Optimal Distributed Generation Location and Sizing using Genetic Algorithms,” IEEE Bucharest Power Tech Conf. 2009, no. 1, pp. 1–6, 2009.
2. M. H. Moradi, M. Abedini, S. M. R. Tousi, and S. M. Hosseinian, “Optimal siting and sizing of renewable energy sources and charging stations simultaneously based on Differential Evolution

algorithm," Int. J. Electr. Power Energy Syst., vol. 73, pp. 1015–1024, 2015.
3. S. Kansal, V. Kumar, and B. Tyagi, "Hybrid approach for optimal placement of multiple DGs of multiple types in distribution networks," Int. J. Electr. Power Energy Syst., vol. 75, pp. 226–235, 2016.
4. M. M. A. Salama, "Distributed generation technologies , definitions and benefits," Electr. Power Syst. Res., vol. 71, no. January, pp. 119–128, 2004.
5. A. Zeinalzadeh, Y. Mohammadi, and M. H. Moradi, "Optimal multi objective placement and sizing of multiple DGs and shunt capacitor banks simultaneously considering load uncertainty via MOPSO approach," Int. J. Electr. Power Energy Syst., vol. 67, pp. 336–349, 2015.
6. N. Ghaffarzadeh and H. Sadeghi, "A new efficient BBO based method for simultaneous placement of inverter-based DG units and capacitors considering harmonic limits," Int. J. Electr. Power Energy Syst., vol. 80, pp. 37–45, 2016.
7. F. Jurado, "A binary SFLA for probabilistic three-phase load flow in unbalanced distribution systems with technical constraints," Int. J. Electr. Power Energy Syst., vol. 48, pp. 48–57, 2013.
8. T. Gözel and M. H. Hocaoglu, "An analytical method for the sizing and siting of distributed generators in radial systems," Electr. Power Syst. Res., vol. 79, no. 6, pp. 912–918, 2009.
9. A. Y. Hatata, G. Osman, and M. M. Aladl, "An optimization method for sizing a solar / wind / battery hybrid power system based on the arti fi cial immune system," Sustain. Energy Technol. Assessments, vol. 27, no. 2018, pp. 83–93, 2018.
10. S. Devi and M. Geethanjali, "Application of Modified Bacterial Foraging Optimization algorithm for optimal placement and sizing of Distributed Generation," Expert Syst. Appl., vol. 41, no. 6, pp. 2772–2781, 2014.
11. N. Khalesi, N. Rezaei, and M. Haghifam, "DG allocation with application of dynamic programming for loss reduction and reliability improvement," Int. J. Electr. Power Energy Syst., vol. 33, no. 2, pp. 288–295, 2011.
12. L. Ioan, M. Abrudean, and D. Bic, "Distributed generation technologies and optimization," Procedia Technol., vol. 12, pp. 687–692, 2014.
13. S. Sultana and P. K. Roy, "Multi-objective quasi-oppositional teaching learning based optimization for optimal location of distributed generator in radial distribution systems," Int. J. Electr. Power Energy Syst., vol. 63, pp. 534–545, 2014.
14. A. K. Singh and S. K. Parida, "Novel sensitivity factors for DG placement based on loss reduction and voltage improvement," Int. J. Electr. Power Energy Syst., vol. 74, pp. 453–456, 2016.
15. N. Jain, S. N. Singh, and S. C. Srivastava, "PSO based placement of multiple wind DGs and capacitors utilizing probabilistic load flow model," Swarm Evol. Comput., vol. 19, pp. 15–24, 2014.
16. H. S. Ramadan, A. F. Bendary, and S. Nagy, "Particle swarm optimization algorithm for capacitor allocation problem in distribution systems with wind turbine generators," Int. J. Electr. Power Energy Syst., vol. 84, pp. 143–152, 2017.
17. A. kumar and K. S. S. Manish kumar, "Optimal Location of WT based Distributed Generation in Pool based Electricity Market using Mixed Integer Non Linear Programming," Mater. Today Proc., vol. 5, no. 1, pp. 445–457, 2018.

****

International Journal of System Dynamics Applications
Volume 11 • Issue 2 • July-December 2022

Models for Efficient Utilization of Resources for Upgrading Android Mobile Technology

Abha Jain, Shaheed Rajguru College of Applied Sciences for Women, India
Ankita Bansal, Netaji Subhas University of Technology, India

ABSTRACT

The need of the customers to be connected to the network at all times has led to the evolution of mobile technology. Operating systems play a vitol role when we talk of technology. Nowadays, Android is one of the popularly used operating system in mobile phones. The authors have analysed three stable versions of Android, 6.0, 7.0, and 8.0. Incorporating a change in the version after it is released requires a lot of rework and thus huge amount of costs are incurred. In this paper, the aim is to reduce this rework by identifying certain parts of a version during early phase of development which need careful attention. Machine learning prediction models are developed to identify the parts which are more prone to changes. The accuracy of such models should be high as the developers heavily rely on them. The high dimensionality of the dataset may hamper the accuracy of the models. Thus, the authors explore four dimensionality reduction techniques, which are unexplored in the field of network and communication. The results concluded that the accuracy improves after reducing the features.

KEYWORDS

Android Operating System, Dimensionality Reduction, Feature Selection, Machine Learning, Mobile Technology, Model Prediction, Software Changes

INTRODUCTION

In today's rapidly growing industry, it is very essential to build an effective and reliable connectivity to establish an efficient communication among the clients and employees of any business organization. Communication involves sharing of critical information between the user and the organization where security plays a key role in terms of privacy (Auxilia et al., 2020). The customers want to be connected and be able to communicate with any business organization at any time and from anywhere. The demand of being connected 24/7 is only possible due to the evolution of mobile and Internet technology. The availability of newer operating systems in the competing market plays a significant role in improving the mobile technology. Among many other characteristics and applications; one of the distinguishing characteristics offered by mobile operating systems is that the users can get connected to the internet using the wireless service provider of their smart phones which is cost effective as the mobile system is completely wireless leading to benefits like saving of money and space as compared to wired sensor network (Elfouly et al., 2017). Various types of operating systems are available in the market, the most popular being Android. Android has seen various versions starting from version 1.0, then1.5, 1.6, 2.1, 2.2, 2.3, 3.0and so on. Each version has the improvements over the previous one and thus, it is always advisable to go with the latest version. The upgradation of a

DOI: 10.4018/IJSDA.20220701.oa2

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

version to a higher version may be due to a number of reasons, such as identification of some bug in the previous version, change in the customers' demands, change in technology in the market, etc. We can also see that change is the need of the hour and we as humans must adapt to those changes. However, seeing from the end of developer's site, we should understand that incorporating a change due to any of the above listed reason resulting in a new version is not at all an easy task. It comes with lots of difficulty and requires huge amount of resources in terms of time, money and manpower. To elaborate on this, let us understand that the development of a software goes through certain stages before it can be deployed. Incorporating a change in any part of the software (due to any of the reasons stated above) may need widespread changes in different parts of the software and thus lots of rework is required (Sharma et al. 2014). Cost and effort of this rework significantly increases with the stages in software development lifecycle (Boehm and Basili 2001).

The authors in this paper aim to work in the direction of reducing the rework and thus, saving of cost and other resources. For this, the three stable versions of Android, viz Android 6.0 (Marshmallow), Android 7.0 (Nougat) and Android 8.0 (Oreo) are analysed. Each version is fairly large in size consisting of a large number of classes. Due to the availability of limited resources, the developers fail to pay equal attention to all the classes, leading to poor quality software. Thus, the main idea revolves around identification of those classes which are more vulnerable to changes in the next software update. The authors have constructed the prediction models which can be used by the developers/designers in the early phases of software development to identify the classes which need focussed attention. Next, we discuss about the correctness or accuracy of the constructed prediction model. Since the developers are relying on these models for identifying the classes, it is very important that the model should be as accurate as possible. The high dimensionality of the dataset is one of the hindrance which may hamper the accuracy of the models. In addition to this, the high index of features (high dimensionality) makes the computation of data an expensive and tedious task (Rattanawadee and Srivihok, 2015). Dimensionality Reduction refers to the process of reducing the number of dimensions of a given data set. This leads to a reduction in the number of variables and utilization of a group of prime variables.

In this paper, the authors have used feature selection algorithms for selecting a subset of relevant features from a given set of features such that they would yield the most optimum results while building an effective and efficient predictive model (Padmaja and Vishnuvardhan, 2016).The literature shows the wide use of traditional statistical method known as regression analysis to extract the useful features. In this study, univariate Logistic Regression (LR) is used to find the effect of each independent variable with the dependent variable. Thereafter, multivariate LR is also used for constructing the model. In addition to the regression analysis which is a statistical approach, there are much newer and popularly used feature selection techniques broadly classified under the three approaches: filter approach, wrapper approach and embedded approach (Bachu and Anuradha 2019). While filter approach focuses on the data instead of the algorithm used for mining it and gathering the relevant information by analysing the nature of the data, the wrapper approach focuses on the applicability or pertinence of each feature and the optimality of the solution thus obtained (Bolón-Canedo et al. 2014). The embedded approach, on the other hand, is more focused and aims towards optimizing the model for a particular training algorithm. In this research, the authors have analysed four sequential search techniques viz. Sequential Forward Selection (SFS), Sequential Backward Selection (SBS), Sequential Forward Floating Selection (SFFS) and Sequential Backward Floating Selection (SBFS) that are employed by wrapper methods for feature selection. The authors found that the usage of these sequential search techniques in different domains such as gene selection, big data classification, pattern recognition and image recognition (Pudil et al. 1993 and Peralta et al. 2015) have produced promising results. However, these techniques are unexplored in the field of networks and communication. Thus, this motivated the authors to explore these search techniques to reduce the features of popular mobile technology, Android.

To determine the efficiency of these sequential search techniques, the prediction models are constructed using three popularly used machine learning classifiers viz. K-Nearest Neighbor (KNN),

Decision Tree (DT) and Random Forest (RF). In other words, after selecting useful features with each of the four sequential search techniques, the dataset is trained and tested using KNN, DT and RF. The authors also observed the relation between the number of features selected and the performance of the searching technique.

The results showed that the maximum accuracy was observed in the mid region where the number of features selected was neither too high nor too low. It was observed that the accuracy increases as the number of features increases from 1 upto the mid - range. There is a slight fluctuation in the accuracy for the mid values and it eventually decreases as the number of features reaches the maximum value. In addition to this, when comparing and analysing the accuracy of the models obtained using multivariate LR and the machine learning models, the authors observed that all the machine learning models outperformed the statistical model.

In this paper, the authors aim to find answers to the following Research Questions (RQs):

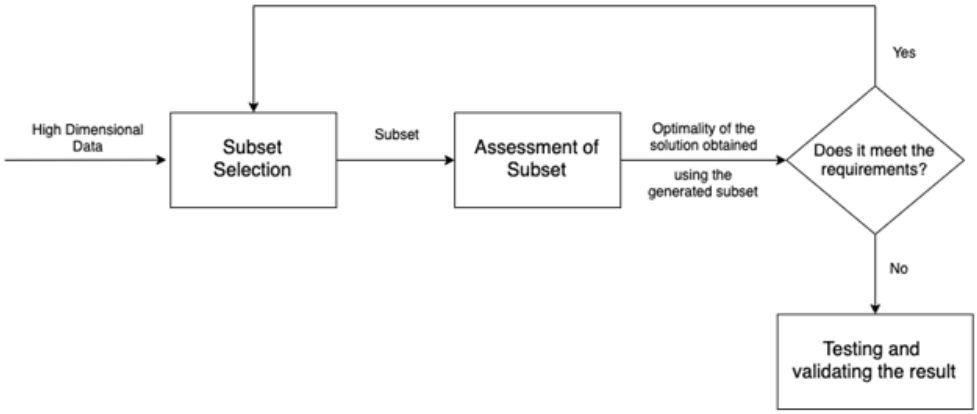
RQ1: How did the sequential search techniques used to reduce the number of features perform in the field of communication, i.e. on the datasets of Android operating system?
RQ2: How did the performance of machine learning classifiers improve when the number of features are reduced?
RQ3: What trend can be observed in the values of accuracy at different number of features selected? Is the highest accuracy achieved at the lowest possible value of number of features?
RQ4: How did multivariate LR model compare with the machine learning models in predicting the change prone classes of Android operating system?

This paper is organized as follows. Following this section, the related work is given which discusses the work done related to feature selection methods. Next section explains the concept and process of feature selection used to select the optimal set of features. Further, it highlights the different kinds of feature selection strategies and various searching techniques which can be used to implement it. After this, the background of the research which includes the details about the empirical data collection and the variables used are explained It also summarizes the empirical data used to validate the results. Next section explains the framework and design of the research. This is followed by the discussion of results. Application of the work and threats to validity are discussed thereafter. Finally, the work is concluded providing important insights.

RELATED WORK

Research shows that the majority of existing work on feature selection does not focus on change prediction. The authors chose to implement various wrapper selection methods for feature reduction to predict change prone classes of open source software as a large number of studies have found that wrapper methods perform best (Kohavi& John, 1997), especially while dealing with lower dimensionality. Wrapper methods like genetic search and sequential forward selection carry out a search over the set of all viable subsets of features, continually calling the induction algorithm as a procedure to assess multiple feature subsets. Work based on feature selection carried out in the past for textual problems turned out to be of great assistance in providing motivation and guidance for this study, which highlights a more substantial variety of metrics. For instance, the authors (Yang & Pedersen, 1997) have examined five feature selection metrics on the common Reuters dataset and OHSUMED. In the papers (Guyon & Elisseeff, 2003 and Liu & Yu, 2005),a thorough study for feature (or variable) selection have been done in the domain of machine learning and statistics. In the paper by Saeys et al. (2007), authors have worked in the domain of bioinformatics by applying feature selection techniques. The work was done in the field of wrapper and filter methods being studied. In the paper by Ma & Huang (2008), the selection of features is done on the basis of sparse regularization. The authors Yang et al. (2013) have proposed an ensemble-based wrapper approach for feature selection

Figure 1. Process of feature selection



from data with highly imbalanced class distribution by creating multiple balanced datasets from the original imbalanced dataset using sampling and then evaluating feature subsets using an ensemble of base classifiers each trained on a balanced dataset. The authors Zhou et al. (2008) presented a Genetic Algorithm (GA) based wrapper method which solves optimization problems using the methods of evolution and is based on survival of the fittest. Their work is based on the classification of hyper spectral data using Support Vector Machine (SVM). The authors (Maldonado and Weber, 2009) had introduced a novel wrapper algorithm for feature selection using SVM with kernel functions. Similar work was also done by the authors Rodriguez-Galiano et al. (2018) and Hu et al. (2015).

FEATURE SELECTION

Feature selection is a technique of dimensionality reduction which being very popular, has been a matter of research in the recent times. It aims at reducing the number of features and providing a subset of the most optimal ones. The process of feature selection process has been highlighted in figure 1. It can be summarized as constituting of the following major steps; creating a subset, assessing the generated subset, checking for the stopping condition and finally testing and validating the result. Following are the advantages of feature selection for dimensionality reduction (Ladha and Deepa, 2011):-

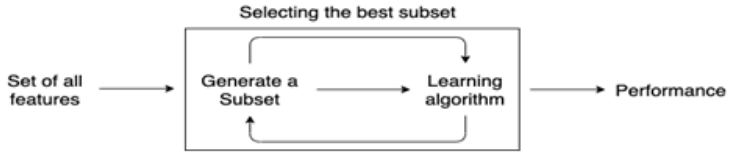
1. Removes the un-important and useless data from the data set.
2. Makes the algorithm more efficient, thereby reducing the time taken by the algorithm to infer results.
3. Overcomes the curse of dimensionality. Curse of dimensionality refers to the difficulties (mainly time and resource constraints) faced in analyzing high dimensional data.
4. Improves the accuracy and performance of the models.

There are three main feature selection methods, namely; filter methods, wrapper methods and embedded methods (Ladha and Deepa 2011). Filter methods are usually implemented in the preconditioning stage. In these methods, the subset of features is picked on the grounds of the correlation value of the features with the output. There are four types of filter methods viz. Linear Discriminant Analysis, Pearson's Correlation, Chi-square and Analysis of Variance. Figure 2 highlights the process followed by filter methods to obtain a feature subset. Wrapper methods train a prediction model by using different subsets of features (Ang et. al, 2015). Based on the performance of the machine learning algorithm, we select the best subset of features from the given dataset. We may choose to include or exclude a feature based on the basis of the performance of a subset. Wrapper

Figure 2. Process followed by filter methods in order to obtain feature subset.



Figure 3. Process followed by wrapper methods to obtain the feature subset.



methods can be classified into two categories, namely; forward selection and backward elimination. In Forward Selection, we start with a void set of features, and keep adding new features in the set with each iteration. On the other hand, in backward elimination, we start a full set of features and keep on eliminating the most useless feature in the entire set with each iteration. Figure 3 shows the process followed by wrapper methods to obtain a feature subset. Embedded methods amalgamate the good attributes of wrapper and filter methods. However, they are complex in implementation. The differences between filter and wrapper methods have been summarized in table 1.Selecting the right feature selection method is a crucial step in determining the efficiency of the algorithm used along with the optimal subset of features to be used. Feature selection speeds up the training of the algorithm used. It also lowers the dimensionality of the dataset, thereby reducing the computational cost and time.

Wrapper methods can be implemented by using various search techniques (Kumari et. al, 2012) which can be categorized as exponential search, random search and sequential search. Exponential search, which may be referred to as total search or complete search, is a comprehensive approach. Although the results that it provides are superlative, this approach takes exponential time to execute. Thus, it cannot be used for large datasets or medium size datasets. An example of the same is Brute-Force search which looks for the solution in each subset. On the other hand, random search, as the name suggests, commences by picking features in a random fashion and then progresses with either of the following search techniques. The first one is a two-way searching approach, like simulated annealing and random hill-climbing. The second one uses techniques which have no uniform variation e.g., Genetic Algorithm (GA) and Tabu search. Finally, sequential search is a greedy optimization technique that finds the local optimum solution. Sequential search can be categorized into four types, namely; Sequential Forward Selection (SFS), Sequential Backward Selection (SBS), Sequential Forward Floating Selection (SFFS) and Sequential Backward Floating Selection (SBFS) which have been explained later in the paper.

Table 1. Difference between wrapper and filter methods

PARAMETER	FILTER METHODS	WRAPPER METHODS
Criteria of selection	Correlation between output variables and features	Utility of a subset of features
Time Complexity	These are faster than wrapper methods because they do not require training of mode with each subset.	These are slower than filter methods.
Evaluation Technique	Statistical methods	Cross Validation
Correctness	Might not be able to find the best solution available	Best Subset of features

Table 2. Comparison between different searching techniques

PARAMETER	EXPONENTIAL	RANDOMISED	SEQUENTIAL
Complexity	Exponential	Polynomial	Polynomial
Type of Algorithm	Brute Force	Greedy	Greedy
Implementation	Difficult to implement	Easier than exponential but more difficult than sequential	Easy to implement
Target dataset size	Small datasets	Small/Medium, but might produce incorrect results	Large datasets

Table 2 illustrates the comparison between different searching techniques used by the wrapper methods. It can be seen from table 2 that the sequential search techniques can be implemented in polynomial time as contrast to exponential search techniques. Although the complexity of randomized search techniques is also polynomial, but it is not suitable for large datasets and it is difficult to be implemented as compared to the sequential search techniques. Since both the datasets used in this paper are large in size, the authors choose to work on sequential search techniques. The authors mainly focus on sequential search techniques (SFS, SBS, SFFS, SBFS) used by wrapper methods and explore their working, optimality and performance. It was observed that the above mentioned techniques have been widely used in different domains such as gene selection, big data classification, pattern recognition and image recognition (Pudil et al. 1993 and Peralta et al. 2015) and have produced promising results. This provides a motivation to the authors to explore these techniques in the field of communication and network.

RESEARCH BACKGROUND

This section focuses on the dataset used highlighting the details of the same. Also, the section elaborates on the independent and dependent variables used in the study.

Empirical Data Collection

For empirical data collection, we have used Android operating system. Android is currently one of the most popular operating system being used in mobile phones and tablets. There are number of versions of Android released in the market till date. The first stable version released was Android 2.3, known by the name of Gingerbread and the latest version is Android 10.0. In this study, we have analyzed the three recent stable versions, Android 6.0, known as Marshmallow, Android 7.0, known as Nougat and Android 8.0, known as Oreo. Since Android is an open source dataset, we downloaded the source code of all these versions from https://source.android.com/source/initializing.html. The details of each version which includes their common name, the total number of classes, the number of classes changed (change - prone classes) and the release date are shown in table 3. The total number of classes is the common classes between the two successive versions. For example, table 3 shows that the total number of classes of Android 6.0 is 10,068. This implies that the common classes between Android 6.0 and Android 7.0 are 10,068.

Independent and Dependent Variables

The dataset consists of 44 independent variables and 1 dependent variable. The independent variables are various Object Oriented (OO) metrics used to determine OO relationships like coupling, cohesion, inheritance etc. The independent variables are provided in table 5. The independent variables are discrete in nature. The dependent variable on the other hand is binary in nature and is used to determine whether the class changes in the next version or not. Its value is 0 for a class if it has not changed

Table 3. Details of the dataset used

Version	Commonly known as	Total Number of Classes	Number of Classes Changed	Release Date
Android 6.0	Marshmallow	10,068	2918	October 5, 2015
Android 7.0	Nougat	11,428	3696	August 22, 2016
Android 8.0	Oreo	-	-	August 21, 2017

in the next version and the value is 1 if it has changed in the next version in terms of number of lines added, deleted and modified. This data consisting of independent and the dependent variable is collected with the help of a tool known as Change Report Generator (CRG) developed by one of the authors (Malhotra et al. 2016).

RESEARCH DESIGN

This section presents the research design used for empirical analysis of change prediction. As depicted in figure 4, the work in this study is conducted in three main phases viz. data filtration using regression analysis and wrapper methods in order to achieve optimal subset of features, thereafter model prediction by applying suitable statistical and machine learning algorithms and finally assessing the performance of the predicted models using suitable performance evaluation measure. Each of the phases has been explained in the subsequent sub-sections along with the methods/techniques followed in that phase.

Data Filtration Using Regression Analysis

In this study, the authors have used a statistical approach known as Logistic Regression (LR) to identify the useful features and for model building. LR is used to predict the dependent variable from a set of independent variables (Hosmer and Lameshow 1989). The authors have used LR as the outcome variable is binary or dichotomous (0 or 1). Both univariate and multivariate regression have been used in this study. Univariate logistic regression is used to find the relationship between the dependent variable and each independent variable. It finds whether there is any significant association between them. Multivariate logistic regression analyses which metrics are useful when they are used in combination. It is used to construct a prediction model used for identifying the change prone classes. To construct the multivariate model, metrics can be fed into the model using two stepwise selection methods, which are forward selection and backward elimination (Hosmer and Lameshow 1989). Forward selection examines the variables that are selected one at a time for entry at each step. The backward elimination method includes all the independent variables in the model and the variables are deleted one at a time from the model until the stopping criteria is fulfilled. However, the results of the model obtained using forward selection were poorer (i.e. the values of $R2$ and log-likelihood statistic defined below were low) than the model obtained from the backward elimination procedure. The authors therefore used backward elimination method in this study.

The general multivariate logistic regression formula is as follows:

Prob (X1, X2, …, Xn) =

where g(x) = B0 + B1*X1 + B2* X2 + … + Bn* Xn

'prob' is the probability of a class being change prone

Xi, (1£ i £ n) are independent variables

Data Filtration Using Wrapper Methods

In this sub-section, we discuss about the working of different wrapper methods which we have applied for the process of feature selection. The reduced feature set that is obtained as a result of

application of these wrapper methods is used for developing the prediction models. According to S´anchez-Maro˜no et al. (2007) and Setiono & Liu (1996), wrapper methods explore all possible combinations of the features and find the one that generates the best results for any given machine learning algorithm. Wrapper methods employ greedy search algorithms. As already mentioned, there are four main types of sequential search techniques (SFS, SBS, SFFS, SBFS) used to implement wrapper methods. Description of each of these search techniques has been presented below. The section also discusses the implementation of these techniques in terms of their pseudo code and how the code was implemented in python. The authors have written the programs in python. The classifiers have been implemented using scikit-learn library. This library has in-built classifiers which can be called using functions. The functions take in various parameters for the classifier to be built. The feature selection algorithm has been made using another python library, namely mlxtend. This is an open source python library which provides machine learning and data science utilities to python's computing stack. They have used the functional sequential feature selector implemented in this library. The function takes in parameters through which we determine if the algorithm is floating or not, and whether the algorithm is forward or not. Furthermore, it takes as an input the classifier which we implemented using the scikit-learn library of python.

In order to understand the working of these techniques, let us assume that there are D features in a dataset but we want only K features.

Sequential Forward Selection (SFS)

According to Marcano-Cedeño et al. (2010), begin with a void set of features and add features to the set one by one, measuring accuracy of the machine learning model at each step. The feature giving the maximum accuracy is retained. The process is repeated K times, in order to store K elements in the set. The final set thus obtained gives the maximum possible accuracy with the desired number of features.

Pseudo code of SFS

Process starts with F1 (a set containing no elements) and the method takes in a number of parameters viz. the classifier to be used, the number of features to select (k), the scoring function and the cross validation. Let Xj be a random variable for feature j and Y be the variable that determines the class label (e.g., healthy vs. unhealthy).
Steps:-

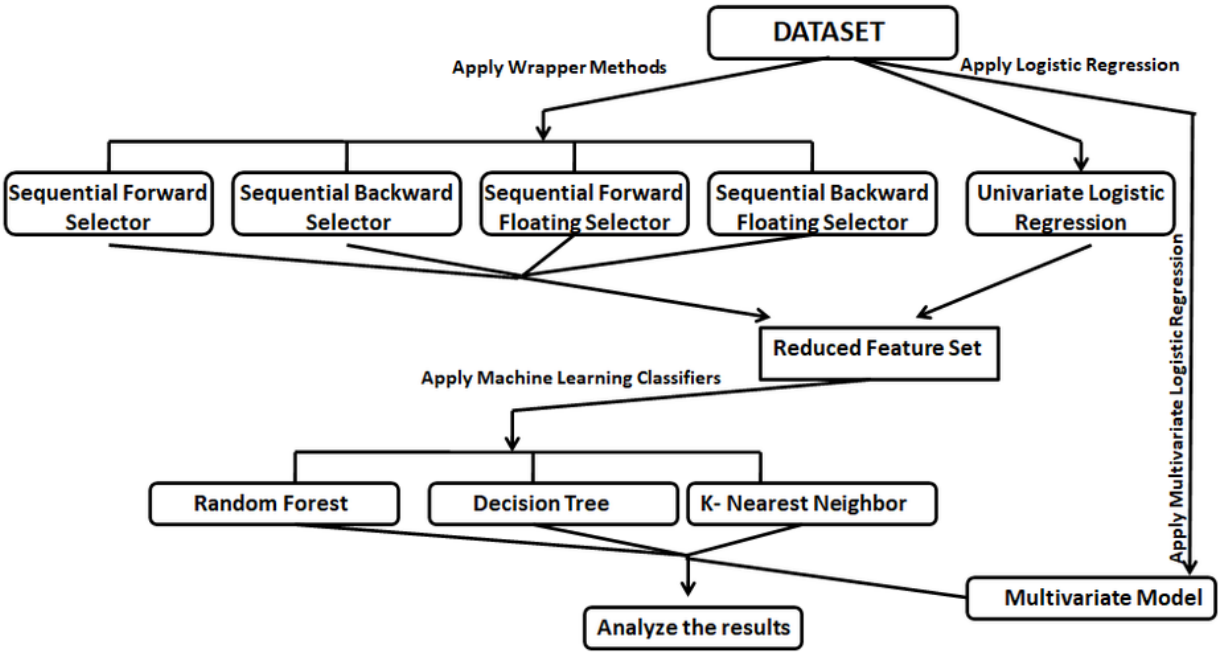
Firstly, a feature X_j is chosen which maximizes the objective function J that takes in the arguments X_j, Y, and F1.

1. The feature Xj that maximizes the objective function is added to F1 and removed from F. This process is repeated until the [REMOVED HYPERLINK FIELD]cardinality of F1 is k.

Sequential Backward Selection (SBS)

Begin with a set containing all the D features. For each of the (D - K) iterations, each feature is removed one by one and the corresponding change in accuracy is noted. The feature with the highest reduction in accuracy is removed. In the end, K features are left which give us the maximum accuracy.

Figure 4. Framework of the study



Pseudo code of SBS

The process starts with F1 (Set of all features), and the method takes in a number of parameters, viz. the classifier to be used, the number of features to select (k), the scoring function and the cross validation. Let Xj be a random variable for feature j and Y be the variable that determines the class label (e.g., healthy vs. unhealthy)
Steps:-
Firstly, feature X_j is chosen which minimizes the objective function J that takes in the arguments X_j, Y, and F1.
1. The feature X_j that minimizes the objective function is removed from F1. This process is repeated until the [REMOVED HYPERLINK FIELD]cardinality of F1 is k.

Sequential Forward Floating Selection (SFFS)

According to Solberg et al (1997), for each step of SFFS, perform the corresponding step of SFS and then find the worst feature in the feature set. Then remove that feature from the feature set if removing it increases the accuracy. Continue removing the worst features as long as the accuracy after removing that feature is greater than the one provided by the SFS set. Then move to the next iteration.

Pseudo code of SFFS

Process starts with F1 (a set containing no elements)
Steps:
Perform SFS.
Find the least significant feature in F1. If it is the feature just added, then keep it and return to step 1. Otherwise, exclude the feature k.
Keep repeating step 2 until we don't have a feature in F1 removing which improves the accuracy.

Sequential Backward Floating Selection (SBFS)

According to Dash et al. (1997) and Wu et al. (2013), for each step of SBFS, perform the corresponding step of SBS and then find the best feature in the feature set. Insert that feature into the feature set if

adding it increases the accuracy. Continue inserting the best features as long as the accuracy after adding that feature is greater than the one provided by the SBS set. Then move to the next iteration.

Pseudo code of SBFS

Steps:-

First perform SBS.

Find the most significant feature in remaining features. If it is the feature just removed, then let it be removed and return to step 1.

Keep repeating step 2 until we don't have a feature in remaining adding which improves the accuracy.

Model Development using Machine Learning Classifiers

Machine learning involves predicting and classifying data and to do so, various machine learning models are used in literature. There are a number of machine learning classifiers available in the literature which are widely used in the classification problems pertaining to diverse fields like software defect prediction and fault severity (Panda, 2019, Hussein et al. 2017), breast cancer classification (Majhi, 2018), health insurance claim prediction (Bhardwaj, 2020) etc. Machine learning classifiers have gained huge popularity in the recent years due to their capability in capturing complex nonlinear relationships among variables. Thus, the authors are motivated to explore the three most popularly used machine learning classifiers viz. K Nearest Neighbour, Random Forest and Decision Tree as their usage is minimal in the field of network for predicting the change prone classes of Android operating system. The overview of these machine learning classifiers is presented in table 4 (Singh et al. 2017, Jain & Vailaya 1996 and Corbane et al. 2009).

Machine learning models have certain parameters (also known as hyperparameters) which can be arbitrarily set by the user before the training process according to a given problem. However, it is challenging to know what values to use for the hyperparameters of a given algorithm on a given dataset. Moreover, there are many hyperparameters associated with each machine learning model and more the hyperparameters we need to tune, the slower the tuning process becomes. Not all model hyperparameters are equally important and some hyperparameters have an outsized effect on the behavior, and in turn, the performance of a machine learning algorithm. Therefore, it is desirable to select a minimum subset of model hyperparameters to search or tune.

In this paper, the authors have used the *Random Search* strategy (Bergstraand and Bengio, 2012) for hyperparameter optimization. In this study, the criteria for selecting the right set of hyperparameters is governed by the following two requirements: (1) minimum execution time and (2) maximum accuracy of the model.

In *Random Search*, a grid of hyperparameters is created which consists of some random values of these hyperparameters. The results of performance would be highly optimistic if training and testing are done the same dataset. Thus, in this study, we have used k-cross validation wherein a single dataset is divided into k parts, out of which one part is used for testing and the remaining other parts are used for training the model (Stone 1974). This process is repeated k number of times, so that each of the k part is used for testing once. In this study, the value of k is taken as 10. In addition to obtaining unbiased results using cross validation, it also has an another important advantage when implementing hyperparameter optimization. Since cross validation allows the dataset to be partitioned into training and testing sets, the authors have avoided using the hyperparameters which worked good on training data but not so good with the test data.

The authors implemented *Random Search* using a utility provided by Scikit-learn (Phython) known as *RandomSearchCV*. Using the scikit-learn *best-estimator attribute*, the authors retrieved the set of hyperparameters which gave the best accuracy of the model in minimum execution time.

Table 4. Description of machine learning techniques along with the values of parameters

Machine learning technique	Description
K Nearest Neighbour (KNN)	KNN is a classifying technique which does not make any assumptions on the underlying data distribution. In KNN, the testing data is classified on the basis of its Euclidean distance from the classes. We take the value of K as 4. The number of parallel workers is set to 10 so as to reduce the total computational time.
Random Forest (RF)	In this method, there are multiple individual decision trees that operate as an ensemble, each of which gives a class prediction and the class with the highest number of votes is returned as the model's prediction. The number of decision trees taken in the random forest is 100. The maximum depth of the tree is set to 2. The seed for generating random numbers is set to 58 which is the best possible value so far. The number of parallel workers is set to 10 so as to reduce the total computational time.
Decision Tree (DT)	It classifies the testing data by forming simple *if-else* statements which are deduced from the features of the training data set. The maximum depth of the tree is set to 2. The seed for generating random numbers is set to 58 which is the best possible value so far. The number of parallel workers is set to 10 so as to reduce the total computational time.

These hyperparameter values is provided in table 4. Once the model is trained, testing is performed using these values of hyperparameters obtained during the training process.

EXPERIMENTAL RESULTS

In this section, we explain the results of regression analysis and the model evaluation results obtained after applying ML classifiers on the features obtained from wrapper methods.

Evaluation of results using regression analysis

In this section, the results of univariate and multivariate LR are discussed (due to space constraint, the results are shown for Android 6.0 only). Similar observations were observed for Android 7.0. Table 5 represents the results of univariate analysis in terms of the coefficient (B) and statistical significance(sig.) for each metric. The parameter "sig" tells whether each of the metric is a significant in predicting the dependent variable (change proneness). If the "sig" value of a metric is below or at the significance threshold of 0.01 or 0.05, then the metric is said to be significant in predicting the change prone classes. In this study, the threshold value is considered as 0.05 (significant values are shown in bold in table 5). The coefficient "B" shows the strength of the independent variable. The higher the value, the higher the impact of the independent variable is. The sign of the coefficient tells whether the impact is positive or negative. Table 5 shows that only the metrics ALB, NOC, CDIM, WMC, NPM and RCC are found to be not significant, whereas all the other metrics are found to be significant predictors of change proneness.

Once the impact of each independent variable on the dependent variable is found, multivariate LR is used to determine the combined effect of independent variables on the dependent variable. All metrics are allowed to enter the model. The variables included in the model are shown in table 6. Table 6 shows that 23 metrics are included in the multivariate model. It can be observed from the table that the 'sig.' value of all the variables included in the model is less than 0.05. The sign of the coefficient of AL, ALB, CCB, CDCM, CDIM,CLB,CLCE,CLC,CS,ME,DIT,SCmetrics is negative, though it was positive in the univariate analysis. This is due to suppressor relationships among independent variables commonly observed in multivariate logistic regression analysis (Briand et al. 2000).

Table 7 shows the confusion matrix for this multivariate model which can be used to evaluate the performance of the model. To evaluate the performance of the models (LR and machine learning),

Table 5. Result analysis of univariate regression

S.No.	Metric	B	Sig.	S.No.	Metric	B	Sig.
1	AC (Average Cyclomatic)	.211	**0.000**	23	CL (Count Line)	.001	**0.000**
2	ACM (Average Cyclomatic Modified)	.291	**0.000**	24	CLB (Count Line Blank)	.002	**0.000**
3	ACS (Average Cyclomatic Strict)	.204	**0.000**	25	LOC (Lines of Code)	.001	**0.000**
4	AE (Average Essential)	.181	**0.000**	26	CLCD (Count Line Code Declared)	.004	**0.000**
5	AL (Average Line)	.042	**0.000**	27	CLCE (Count Line Code Executed)	.004	**0.000**
6	ALB (Average Line Blank)	.026	0.099	28	CLC (Count Line Comment)	.006	**0.000**
7	ALC (Average Line Code)	.050	**0.000**	29	(CS) Count Semicolon	.004	**0.000**
8	ALCo (Average Line Comment)	.078	**0.000**	30	CS (Count Statement)	.002	**0.000**
9	CCB (Count Class Base)	.195	**0.000**	31	CSD (Count Statement Declared)	.003	**0.000**
10	CBO (Coupling Between Objects)	.110	**0.000**	32	CSE (Count Statement Executed)	.004	**0.000**
11	NOC (Number Of Children)	.000	0.697	33	MC (Maximum Cyclomatic)	.112	**0.000**
12	CDCM (Count Declared Class Method)	.060	**0.000**	34	MCM (Maximum Cyclomatic Modified)	.145	**0.000**
13	CDCV (Count Declared Class Variable)	.045	**0.000**	35	MCS (Maximum Cyclomatic Strict)	.097	**0.000**
14	CDIM (Count Declared Instance Method)	.000	0.754	36	ME (Maximum Essential)	.177	**0.000**
15	CDF (Count Declared Function)	-.896	**0.000**	37	DIT (Depth of Inheritance Tree)	-.108	**0.000**
16	CDIV (Count Declared Instance Variable)	.129	**0.000**	38	MN (Maximum Nesting)	.544	**0.000**
17	WMC (Weighted Methods per Class)	.042	0.327	39	LCOM (Lack of Cohesion of Methods)	.019	**0.000**
18	RFC (Response For a Class)	.000	**0.007**	40	RCC (Ratio Comment to Code)	.009	0.422
19	CDMD (Count Declared Method Default)	.207	**0.000**	41	SC (Sum Cyclomatic)	.002	**0.000**
20	NPRM (Number of Private Methods)	.171	**0.000**	42	SCM (Sum Cyclomatic Modified)	.002	**0.000**
21	NPROM (Number of Protected Methods)	.146	**0.000**	43	SCS (Sum Cyclomatic Strict)	.002	**0.000**
22	NPM (Number of Public Methods)	.000	0.114	44	SE (Sum Essential)	.001	**0.000**

Table 6. Metrics included in the multivariate regression model

S.No.	Metrics included in the model	B	Sig.	S.No.	Metrics included in the model	B	Sig.
1	SE	.025	.001	12	CL	.011	.000
2	AL	-.026	.010	13	CLB	-.027	.000
3	ALB	-.160	.000	14	CLCE	-.007	.000
4	ALC	.052	.000	15	CLC	-.008	.000
5	CCB	-.099	.032	16	CS	-.006	.001
6	CBO	.081	.000	17	CSD	.007	.004
7	CDCM	-.026	.001	18	MC	.033	.000
8	CDIM	-.020	.003	19	ME	-.056	.000
9	CDIV	.024	.000	20	DIT	-.231	.000
10	RFC	.001	.000	21	LCOM	.004	.000
11	CDMD	.065	.000	22	SC	-.028	.001
				23	SCM	.016	.028

Table 7. Confusion matrix for multivariate logistic regression

		Predicted	
		0 (Not change prone)	1 (Change prone)
Actual	0 (Not change prone)	6730	420
	1 (Change prone)	1824	1094

a popularly used performance measure known as accuracy is employed. Accuracy is the fraction of predictions the model got right. In other words, it is defined as the ratio of correct predictions to the total number of predictions (both correct and incorrect). It can be seen from table 7 that the out of a total of 10068 classes, 7824 classes are correctly predicted. In other words, 6730 classes are correctly predicted to be change prone and 1094 classes are correctly predicted to be not change prone. Thus, the accuracy of the LR model is 77.77% which is good but much less as compared to the accuracy of the models achieved using machine learning classifiers (discussed in coming section).

Evaluation of results using wrapper methods

This section depicts the performance of three classifiers used in the study viz. KNN, RF and DT in terms of accuracy when the features have been reduced using four sequential search techniques of wrapper methods viz. SFS, SBS, SFFS and SBFS.

The authors depict the results with the values obtained for k=1,9,18,27,36,44 where k is the number of features obtained after feature selection. The authors aim to analyze the trend which can be observed in the accuracy obtained on the basis of the number of features selected and thereby studying that in order to get the best possible accuracy, should the number of features selected be high, low or average. Table 8 and 9 show the performance of classifiers when empirical validation has been conducted on Android 6.0 and Android 7.0 dataset respectively. The accuracy results are

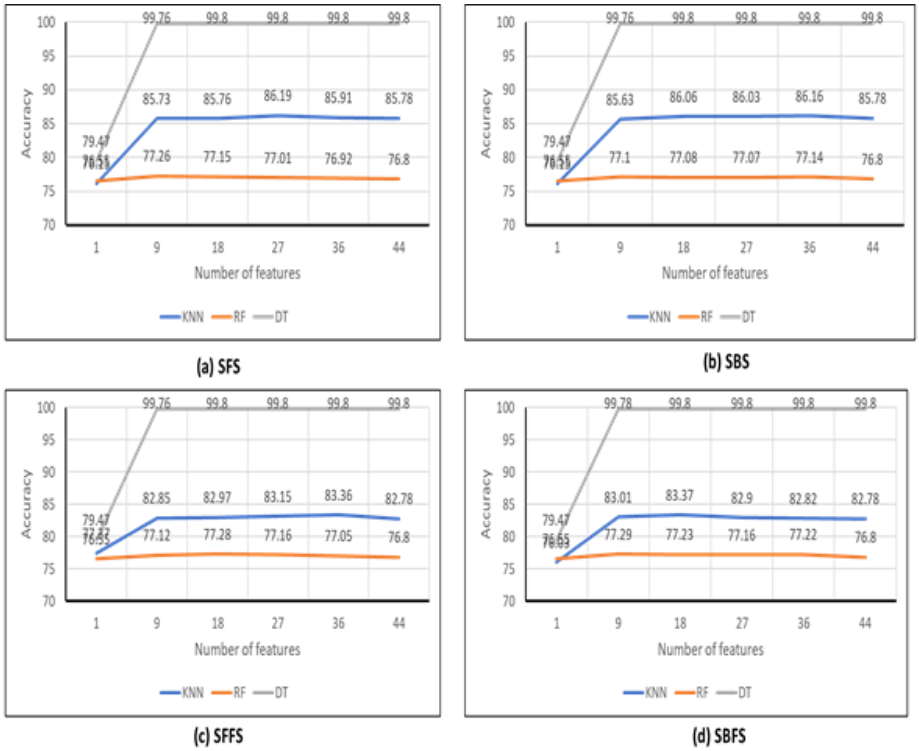
Table 8. Performance of classifiers with respect to different search techniques applied on Android 6.0dataset

SFS				**SBS**			
No. of features	**KNN**	**RF**	**DT**	**No. of features**	**KNN**	**RF**	**DT**
1	76.15	76.55	79.47	1	76.15	76.55	79.47
9	85.73	**77.26**	99.76	9	85.63	77.10	99.76
18	85.76	77.15	**99.80**	18	**86.06**	**77.08**	**99.80**
27	**86.19**	77.01	99.80	27	86.03	77.07	99.80
36	85.91	76.92	99.80	36	86.16	77.14	99.80
44	85.78	76.80	99.80	44	85.78	76.80	99.80
SFFS				**SBFS**			
No. of features	**KNN**	**RF**	**DT**	**No. of features**	**KNN**	**RF**	**DT**
1	77.37	76.55	79.47	1	76.03	76.55	79.47
9	82.85	77.12	99.76	9	83.01	**77.29**	99.78
18	82.97	**77.28**	**99.80**	18	**83.37**	77.23	**99.80**
27	83.15	77.16	99.80	27	82.90	77.16	99.80
36	**83.36**	77.05	99.80	36	82.82	77.22	99.80
44	82.78	76.80	99.80	44	82.78	76.80	99.80

Table 9. Performance of classifiers with respect to different search techniques applied on Android 7.0 dataset

SFS				**SBS**			
No. of features	**KNN**	**RF**	**DT**	**No. of features**	**KNN**	**RF**	**DT**
1	74.69	74.54	77.11	1	72.82	74.54	77.11
9	81.51	**75.28**	99.73	9	84.99	**75.08**	99.73
18	81.53	75.14	**99.78**	18	85.30	74.84	**99.78**
27	**81.70**	75.21	99.78	27	**85.53**	74.88	99.78
36	81.89	74.99	99.78	36	85.57	75.04	99.78
44	81.50	73.96	99.78	44	81.50	73.96	99.78
SFFS				**SBFS**			
No. of features	**KNN**	**RF**	**DT**	**No. of features**	**KNN**	**RF**	**DT**
1	74.69	74.54	77.11	1	74.69	74.54	77.11
9	81.51	**75.34**	99.73	9	81.51	**75.19**	99.73
18	81.53	75.25	**99.78**	18	81.53	75.17	**99.78**
27	81.95	75.03	99.78	27	**81.95**	75.11	99.78
36	**82.14**	75.03	99.78	36	81.76	75.05	99.78
44	81.50	73.96	99.78	44	81.50	73.96	99.78

Figure 5. Accuracy results of classifiers when feature selection of Android 6.0 dataset has been done using (a) SFS, (b) SBS, (c) SFFS, (d) SBFS

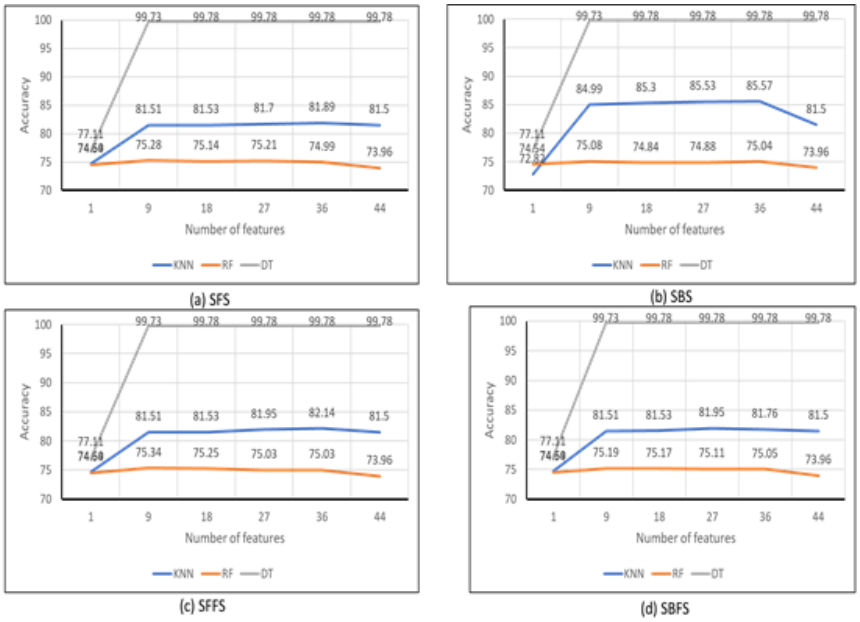


also graphically represented in figures5(a-d) and 6(a-d) depicting the performance of classifiers when SFS, SBS, SFFS and SBFS techniques have been applied on Android 6.0 and Android 7.0 dataset respectively.

Let us first discuss the accuracy achieved by each of the classifiers when SFS technique has been used for feature selection on Android 6.0 dataset. As it can be observed from table 8, DT achieves the highest accuracy values for all the values of k, while RF achieves the lowest values of accuracy. The accuracy in the case of each of the classifier increases as the value of k increases for lower values of k, however, it decreases when k approaches the maximum value. The highest accuracy value for each classifier is shown in bold. Analyzing the values of k at these highest accuracies show that in none of the cases the highest accuracy is attained at k=1 or k=44. This shows that when features are reduced, the accuracy increases but we have to carefully decide the minimum number of features. The accuracy value corresponding to DT begins at 79.47% for k=1, keeps increasing till k=18, attaining the value of 99.8% and then remains constant. The accuracy value for KNN begins at 76.15% for k=1 and increases until k=27, reaching the value 86.19% and then starts decreasing. The accuracy value observed for KNN is 85.78%when k=44, lower than the accuracy value corresponding to DT but higher than that of RF. The accuracy value with respect to RF begins at 76.55% for k=1, a value that is lower than the corresponding value of accuracy for DT but higher than KNN. The accuracy increases till k=9, reaching as high as 77.26% and then decreases continuously to stop at 76.80% for k=44. It can be observed that the most optimal results in general are obtained when the value of k= 18 with DT chosen as the classifier. This observation is consistent with other feature selection techniques viz. SBS, SFFS and SBFS too.

For Android 7.0, similar observations can be inferred. The highest accuracies are obtained at k=9 for RF and k=18 for DT using all the feature selection techniques. When KNN is used, highest accuracies are obtained at k=27 for all the feature selection techniques except SFFS which shows highest accuracy at k=36. DT shows the highest accuracy of 99.78% at k=18 amongst all the

Figure 6. Accuracy results of classifiers when feature selection of Android 7.0 dataset has been done using (a) SFS, (b) SBS, (c) SFFS, (d) SBFS



classifier with all the feature reduction techniques. As discussed above, same conclusion is drawn from the results on Android 6.0. Thus, the authors recommend the use of DT for model prediction after reducing the number of features.

APPLICATION OF THE WORK

The usage of software in every field of Network and Communication has become an integral part of everyone's lives. Due to multiple reasons such as ever changing demands of the customers, change in technology in the market, identification of some bug etc. lead to upgradation of a software from one version to the next. When software developers work on upgrading the version, lots of effort is required leading to utilization of large amount of resources. The results of this work will be of interest to researchers as well as practitioners from industry in reducing this effort in terms of time, money and manpower required in development of software, thus leading to the delivery of software with better quality. This is done by identifying those classes amongst other classes which may change during later phases of development and thus, need careful attention by the developers, designers and testers. Such classes ae termed as 'change - prone' classes. The authors in this paper have developed various machine learning models which can be used to predict change-prone classes. Timely identification of such classes would be of great benefit to the software developers from industry as these classes play a critical role in design and architecture of the system. The architecture of the system once built acts as the scaffolding in which the functionality of the system is delivered, thus ensuring that system delivered meets the customer's functional expectations and needs. Once the change - prone classes are identified during the early phases, the architecture of the system can be altered easily. For example, if a class 'A' is predicted to be change - prone, using the software metrics we can check its coupling, cohesion, inheritance etc. Accordingly we can modify the architecture in the design phases such that the values of these software metrics fall within the required range. In other words, certain corrective measures can be taken to ensure the software delivered is error-free which is actually significant when different versions of software are released frequently. The measurements derived from design can be used as benchmark in organizations, to assess the quality of software products. Releasing software

of good quality leads to happy and satisfied customer. This leads in increasing the reputation and status of the software organization (in which the software is developed) in the market. Thus, customer satisfaction which is of utmost importance in today's scenario is met. Also, design of the classes of future release can be re-checked and better designs can be suggested.

THREATS TO VALIDITY

The empirical validation in this work has certain limitations which may adversely affect the validity of the results. These limitations are discussed in terms of four threats to validity, viz. construct validity, internal validity, external validity and conclusion validity.

1. Construct Validity

This type of validity is one of the most important threats to validity. It is defined as the extent to which the variables (independent and dependent variables) and the performance parameters precisely measure the concept they intend to measure (Dean and Voss 1999, Zhou et al. 2009). This threat can be due to the improper collection of the dependent variable and the independent variables. There have been studies in research which have determined the accuracy of some of the OO metrics employed in this study(Briand et al. 1998, 1999, 2000). As a result, this threat is reduced in the study. The dependent and the independent variables in this study are collected using the tool known as CRG (Malhotra et al. 2016) developed by the authors themselves. Due to this, the exact steps or procedure to collect the dependent and the independent variables is known, which is important to provide for an accurate assessment of their construct validity. Hence, the threat due to improper collection of the variables is also reduced in this study. However, the authors have not taken into consideration the type of change (corrective, adaptive, perfective or preventive) a class may go through which may pose a threat on the evaluation of the results. In the future work, this threat can be reduced by considering the type of change as well.

2. Internal Validity

This validity is defined as the degree to which conclusions can be drawn about the causal effect of independent variable on the dependent variable" (Zhou et al. 2009). In this work, independent variables used are a set of OO metrics measuring different concepts of OO paradigm. These independent variables are not related to each other in any way. All these metrics together determine the value of the dependent variable. It is not possible to determine the causal effect of each independent variable on the dependent variable. In other words, goal of this study is to develop prediction models that will identify change - prone classes rather than to discover the cause-effect relationships. Thus, the threat to internal validity does not exist in the study.

3. External Validity

This validity deals with the generalization of the results obtained by the study. In other words, it concerns itself with finding out whether the results produced by the study are applicable in different domains or can be replicated in different scenarios for which the results are not evaluated (Harrison et al. 2000). For this, all such information which is required for replicating the study should be transparently available in the study. For example, the information about the availability of the dataset, nature of the dataset, details of the approaches used in the paper and their default or tuned parameters etc. should be clearly stated in the study. Since, in the present work, Android dataset used is an open source dataset, the source codes of all the versions can be downloaded from https://source.android.

com/source/initializing.html. In addition to this, the authors have provided complete details of the three versions that are analysed in this study in table 3. Moreover, for empirical validation, the authors have written the programs in python. The classifiers have been implemented using scikit-learn library. This library has in-built classifiers which can be called using functions. The functions take in various parameters for the classifier to be built. The feature selection algorithm has been made using another python library, namely mlxtend. The pseudo code of each of the feature selection algorithm is provided and the parameters used for the machine learning classifiers are also mentioned in table 4. Using all this information, the results can be replicated and generalized across different datasets. In this study, the authors have worked on different releases of software that large in size with the aim to draw meaningful conclusions. However, we cannot consider it as a complete generalization wherein the results could be applied universally. This is so because we are dealing with multiple versions of a single software as of now. In the future, we will be comparing the results of different datasets across different projects having diverse characteristics.

4. Conclusion Validity

This threat includes all those threats that may affect the conclusion of the study. In other words, all the threats which may lead to improper results or conclusions of the study are called as conclusion validity threats (Malhotra 2016). The authors in this study have not performed the statistical evaluation of the results using statistical tests. Thus, this leads to a conclusion validity threat.

CONCLUSION AND FUTURE WORK

The upcoming technologies in mobile phones have evolved rapidly over the last some years. Nowadays, one of the essential needs of human beings is that of mobile phones. It helps everyone to be connected to each other all the time and all the places. Amongst many other factors playing important role in improving mobile's technology, operating system is one of the important factor too. Android, Symbian, iOS, Palm OS etc. are some of the commonly used operating systems of mobile integrated with multiple user-friendly features. The most popularly and widely used mobile operating system is Android. Android has seen multiple versions, the most recent one being Android 10.0 (officially known as Android 10).Advancing from one version to a newer version requires a lot of effort in terms of resources utilized. We know that the resources such as time, money and manpower are very limited and thus, judicious use of such resources is very essential. In this paper, the broad objective is to reduce this effort, thus leading to saving of resources. For doing this, machine learning models are constructed which can be used during the initial phases of software development for identifying certain parts of software which have a higher tendency to change in the later stages. Once such parts (classes) are known, careful attention is required to be given to them in order to avoid wastage of essential resources. The accuracy of such models should be high and high dimensionality play a very important role in the accuracy of the models. The authors in this paper are working on feature selection methods to reduce the dimensionality of the data.

In other words, feature selection is an important method that can be used for dimensionality reduction of multi-dimensional data that we often encounter in real life situations. Feature selection can be achieved via wrapper methods, filter methods or embedded methods but the authors choose wrapper methods as they analyze the utility of a subset of features and return the best possible subset. Feature selection can also be viewed in terms of searching techniques like sequential search, exponential search and random search. However, the authors implemented sequential search techniques as they are greedy optimization techniques that find the local optimum solution in polynomial time complexity and are easy to implement, even with large datasets. Authors focused on the implementation of the four major sequential feature selection techniques which are Sequential Forward Selection (SFS), Sequential

Backward Selection (SBS), Sequential Forward Floating Selection (SFFS) and Sequential Backward Floating Selection (SBFS). In addition to the wrapper methods for selecting useful features, the authors have also used statistical approach of regression analysis to extract the useful features. Univariate and multivariate Logistic Regression have been used to extract useful features and constructing the prediction model respectively. Empirical validation was conducted against three stable versions of Android operating system, Android 6.0, Android 7.0 and Android 8.0. The dataset comprised of 44 independent variables. The aim is to reduce the number of features and select the most relevant ones from this set. In order to achieve the above objective, prediction models are constructed using three machine learning classifiers viz. K-Nearest Neighbor (KNN), Decision Tree (DT) and Random Forest (RF). These models are used to identify those change prone classes of the software in early phases of software development, leading to optimum utilization of limited resources.

Following are the important insights gathered from the results:

- DT classifier outperformed the remaining two classifiers (KNN and RF) depicting high values of accuracy for each of the sequential feature selector used.
- With respect to KNN classifier, it was observed that SBS is the best sequential selector. On the other hand, with respect to RF classifier, SFFS sequential classifier is the best.
- Highest value of accuracy was achieved when the number of features selected was 9 or 18 (in general). Therefore, it can be inferred that either too high(44) or too low(1) number of features selected leads to decrease in accuracy. So, we must choose the number of features which is neither too high nor too low.
- The machine learning models (built using KNN, DT and RF) outperformed the statistical model (built using multivariate Logistic Regression) by depicting higher values of accuracy. Thus, the authors suggest the use of machine learning models for predicting the change prone classes of Android operating system.

As future work, the authors plan to work on feature selection methods based on other searching techniques as well, i.e. exponential search and random search so that a comparative analysis can be done. For the purpose of empirical validation, the study can be extended on different datasets belong to diverse domains. In addition to this, more machine learning techniques could be used so that a fair and exhaustive evaluation is possible.

REFERENCES

Ang, J. C., Mirzal, A., Haron, H., & Hamed, H. N. A. (2016). Supervised, Unsupervised and Semisupervised Feature Selection: A Review on Gene Selection. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *13*(5), 971–989. doi:10.1109/TCBB.2015.2478454 PMID:26390495

Auxilia, M., Raja, K., & Kannan, K. (2020). Cloud-Based Access Control Framework for Effective Role Provisioning in Business Application. *International Journal of System Dynamics Applications*, *9*(1), 1–18. doi:10.4018/IJSDA.2020010104

Bachu, V., & Anuradha, J. (2019). A Review of Feature Selection and Its Methods. *Cybernetics and Information Technologies*, *19*(1), 3–26. doi:10.2478/cait-2019-0001

Bergstra, J., & Bengio, Y. (2012). Random Search for Hyper-Parameter Optimization. *Journal of Machine Learning Research*, *13*, 281–305.

Bhardwaj, A. (2020). Health Insurance Claim Prediction Using Artificial Neural Networks. *International Journal of System Dynamics Applications*, *9*(3), 1–18.

Boehm, B., & Basili, V. R. (2001). Software defect reduction top 10 list. *Computer*, *34*(1), 135–137. doi:10.1109/2.962984

Bolón-Canedo, V., Sánchez-Maroño, N., & Alonso-Betanzos, A. (2014). Data classification using an ensemble of filters. *Neurocomputing*, *135*, 13–20. doi:10.1016/j.neucom.2013.03.067

Briand, L., Daly, J., & Wust, J. (1998). A unified framework for cohesion measurement in object-oriented systems. *Empirical Software Engineering*, *3*(1), 65–117. doi:10.1023/A:1009783721306

Briand, L., Daly, J., & Wust, J. (1999). A unified framework for coupling measurement in object-oriented systems. *IEEE Transactions on Software Engineering*, *25*(1), 91–121. doi:10.1109/32.748920

Briand, L., Wust, J., Daly, J. W., & Victor Porter, D. (2000). Exploring the relationship between design measures and software quality in object-oriented Systems. *Journal of Systems and Software*, *51*(3), 245–273. doi:10.1016/S0164-1212(99)00102-8

Corbane, C., Baghdadi, N., Descombes, X., Wilson, G. J., Villeneuve, N., & Petit, M. (2009). Comparative Study on the Performance of Multiparameter SAR Data for Operational Urban Areas Extraction Using Textural Features. *Geoscience and Remote Sensing Letters, IEEE*, *6*(11), 728–732. doi:10.1109/LGRS.2009.2024225

Dash, M., Liu, H., & Yao, J. (1997). Dimensionality Reduction of Unsupervised Data. *Proceedings Ninth IEEE International Conference on Tools with Artificial Intelligence*. doi:10.1109/TAI.1997.632300

Dean, A., & Voss, D. (1999). *Design and analysis of experiments*. Springer. doi:10.1007/b97673

Elfouly, F. H., Ramadan, R. A., Mahmoud, M. I., & Dessouky, M. I. (2017). Efficient Data Reporting in a Multi-Object Tracking Using WSNs. *International Journal of System Dynamics Applications*, *6*(1), 1–20. doi:10.4018/IJSDA.2017010103

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, *3*, 1157–1182.

Hall, M. A., & Smith, L. A. (1999). Feature Selection for Machine Learning: Comparing a Correlation-based Filter Approach to the Wrapper. *Proceedings of the Twelfth International Florida Artificial Intelligence Research Society Conference*, 235-239.

Harrison, R., Counsell, S., & Nithi, R. (2000). Experimental Assessment of the Effect of Inheritance on the Maintainability of Object-Oriented Systems. *Journal of Systems and Software*, *52*(2-3), 173–179. doi:10.1016/S0164-1212(99)00144-2

Hosmer, D., & Lemeshow, S. (1989). *Applied logistic regression*. Wiley.

Hu, Z., Bao, Y., Xiong, T., & Chiong, R. (2015). Hybrid filter–wrapper feature selection for short-term load forecasting. *Engineering Applications of Artificial Intelligence, 40,* 17–27.

Hussein, H. A. T., Ammar, M. E., & Moustafa Hassan, M. A. (2017). Three Phase Induction Motor's Stator Turns Fault Analysis Based on Artificial Intelligence. *International Journal of System Dynamics Applications*, *6*(1), 1–19. doi:10.4018/IJSDA.2017070101

Jain, K., & Vailaya, A. (1996). Image Retrieval Using Color and Shape. *Pattern Recognition*, *29*(8), 1233–1244. doi:10.1016/0031-3203(95)00160-3

Jain, A., & Zonkar, D. (1997). Feature Selection: Application, Evaluation and small sample performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *19*(2), 153–158. doi:10.1109/34.574797

Kagdi, H., & Maletic, J.I. (2016). Software-Change Prediction: Estimated+Actual. *International Journal of Computer Applications in Technology, 54*(4), 240-256.

Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. *Artificial Intelligence*, *97*(1-2), 273–324. doi:10.1016/S0004-3702(97)00043-X

Kumar, V., & Minz, S. (2015). Feature Selection: A literature Review. *Proceedings of the Third International Symposium on Women in Computing and Informatics, 31-37.*

Kumari, A., Tripathi, R., Pal, M., & Chakraborty, S. (2012). Linear Search versus Binary Search: A statistical comparison for binomial inputs. *International Journal of Computer Science Engineering and Applications*, *2*(2), 29–39.

Ladha, L., & Deepa, T. (2011). Feature selection methods and algorithms. *International Journal on Computer Science and Engineering*, *3*(5), 1787–1797.

Liu, H., & Yu, L. (2005). Toward integrating feature selection algorithms for classiðcation and clustering. *IEEE Transactions on Knowledge and Data Engineering*, *17*(4), 491–502. doi:10.1109/TKDE.2005.66

Ma, S., & Huang, J. (2008). Penalized feature selection and classiðcation in bioinformatics. *Briefings in Bioinformatics*, *9*(5), 392–403. doi:10.1093/bib/bbn027 PMID:18562478

Majhi, S. K. (2018). An Efficient Feed Foreword Network Model with Sine Cosine Algorithm for Breast Cancer Classification. *International Journal of System Dynamics Applications*, *7*(2), 1–14. doi:10.4018/IJSDA.2018040101

Malhotra, R. (2016). *Empirical research in software engineering: concepts, analysis, and applications*. CRC Press, Taylor & Francis. doi:10.1201/b19292

Malhotra, R., Bansal, A., & Jajoria, S. (2016). An Automated Tool for Generating Change Report from Open-Source Software. *Int. Conference on Advances in Computing, Communications and Informatics (ICACCI)*. doi:10.1109/ICACCI.2016.7732273

Maldonado, S., & Weber, R. (2009). A wrapper method for feature selection using Support Vector Machines. *Information Sciences, 179,* 2208–2217.

Marcano-Cedeño, A., Quintanilla-Domínguez, J., Cortina-Januchs, M. G., & Andina, D. (2010). Feature Selection Using Sequential Forward Selection and classification applying Artificial Metaplasticity Neural Network. *36th Annual Conference on IEEE Industrial Electronics Society (IECON).* doi:10.1109/IECON.2010.5675075

Padmaja, D. L., & Vishnuvardhan, B. (2016). Comparative Study of Feature Subset Selection Methods for Dimensionality Reduction on Scientific Data. *IEEE 6th International Conference on Advanced Computing (IACC).*

Panda, M. (2019). Software Defect Prediction Using Hybrid Distribution Base Balance Instance Selection and Radial Basis Function Classifier. *International Journal of System Dynamics Applications*, *8*(3), 53–75. doi:10.4018/IJSDA.2019070103

Peralta, D., Rios, S., Ramirez-Gallegos, S., Triguero, I., Benitez, J. M., & Herrera, F. (2015). Evolutionary Feature Selection for Big Data Classification: A MapReduce Approach. *Mathematical Problems in Engineering*, *2015*, 1–11. doi:10.1155/2015/246139

Pudil, P., Novovicova, J., & Kittler, J. (1993). Floating search methods in feature selection. *Pattern Recognition Letters*, *15*(11), 1119–1125. doi:10.1016/0167-8655(94)90127-9

Rattanawadee, P., & Srivihok, A. (2015). Wrapper Feature Subset Selection for Dimension Reduction Based on Ensemble Learning Algorithm. *Procedia Computer Science*, *72*, 162–169. doi:10.1016/j.procs.2015.12.117

Rodriguez-Galiano, V.F., Luque-Espinar, J.A., Chica-Olmo, M., & Mendes, M.P. (2018). Feature selection approaches for predictive modelling of groundwater nitrate pollution: An evaluation of filters, embedded and wrapper methods. *Science of the Total Environment, 624,* 661–672.

Saeys, Y., Inza, I., & Larraaga, P. (2007). A review of feature selection techniques in bioinformatics. *Bioinformatics (Oxford, England)*, *23*(19), 2507–2517. doi:10.1093/bioinformatics/btm344 PMID:17720704

S'anchez-Maro˜no, N., Alonso-Betanzos, A., & Tombilla-Sanrom'an, M. (2007). Filter Methods for Feature Selection- A Comparative Study. In H. Yin, P. Tino, E. Corchado, W. Byrne, & X. Yao (Eds.), Lecture Notes in Computer Science: Vol. 4881. *Intelligent Data Engineering and Automated Learning - IDEAL 2007* (pp. 178–187). Springer. doi:10.1007/978-3-540-77226-2_19

Setiono, R., & Liu, H. (1996). A probabilistic approach to feature selection - a filter solution. *ICML'96 Proceedings of the Thirteenth International Conference on International Conference on Machine Learning Pages,* 319-327.

Sharma, V. S., Ramnani, R. R., & Sengupta, S. (2014). A framework for identifying and analyzing non-functional requirements from text. *Proceedings of the 4th International Workshop on Twin Peaks of Requirements and Architecture*. doi:10.1145/2593861.2593862

Singh, A., Halgamuge, M. N., & Lakshmiganthan, R. (2017). Impact of Different Data Types on Classifier Performance of Random Forest, Naïve Bayes, and K-Nearest Neighbors Algorithms. *International Journal of Advanced Computer Science and Applications*, *8*(12). Advance online publication. doi:10.14569/IJACSA.2017.081201

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Society of Arts*, *36*(2), 111–114. doi:10.1111/j.2517-6161.1974.tb00994.x

Yang, P., Liu, W., Zhou, B. B., Chawla, S., & Zomaya, A. Y. (2013). Ensemble-Based Wrapper Methods for Feature Selection and Class Imbalance Learning. In J. Pei, V. S. Tseng, L. Cao, H. Motoda, & G. Xu (Eds.), Lecture Notes in Computer Science: Vol. 7818. *Advances in Knowledge Discovery and Data Mining. PAKDD 2013*. Springer. doi:10.1007/978-3-642-37453-1_45

Yang, Y., & Pedersen, J. O. (1997). A Comparative Study on Feature Selection in Text Categorization. *ICML '97 Proceedings of the Fourteenth International Conference on Machine Learning Pages,* 412-420.

Wu, S., Hu, Y., Wang, W., Feng, X., & Shu, W. (2013). Application of global optimization methods for feature selection and machine learning. *Mathematical Problems in Engineering*, ●●●, 1–9.

Zhou, Y., Leung, H., & Xu, B. (2009). Examining the Potentially Confounding Effect of Class Size on the Associations between Object-Oriented Metrics and Change-Proneness. *IEEE Transactions on Software Engineering*, *35*(5), 607–623. doi:10.1109/TSE.2009.32

Zhuo, L., Zheng, J., Li, X., Wang, F., Ai, B., & Qian, J. (2008). A genetic algorithm based wrapper feature selection method for classification of hyperspectral images using support vector machine. *Proc. SPIE 7147, Geoinformatics and Joint Conference on GIS and Built Environment: Classification of Remote Sensing Images,* 71471. doi:10.1117/12.813256

J. A. Jain is an Assistant Professor at the Department of Computer Science Engineering, Shaheed Rajguru College of Applied Sciences for Women, Delhi University, India. Prior to joining the college, she worked as full-time research scholar and received a doctoral research fellowship from Delhi Technological University (formerly Delhi College of Engineering). She received her master's and doctorate degree in software engineering from Delhi Technological University. Her research interests are IoT, networks and communication, data mining, software quality, and statistical and machine learning models. She has published papers in International journals and conferences.

A. Bansal is an Assistant Professor at the Department of Information Technology, Netaji Subhas University of Technology (formerly known as Netaji Subhas Institute of Technology), Delhi, India. Prior to joining the university, Ankita Bansal worked as full-time research scholar at Delhi Technological University (formerly Delhi College of Engineering), Delhi, India. She received her master's and doctoral degree in computer science from Delhi College of Engineering. Her research interests are network models, communication technology, IoT, software quality, soft computing, database management, machine learning and meta heuristic models. She has published papers in international journals and conferences.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3104828, IEEE Transactions on Power Delivery

Modified CCF based Shunt Active Power Filter Operation with Dead-Band Elimination for Effective Harmonic and Unbalance Compensation in 3-phase 3-wire System

Alok Kumar Dubey, *Student Member, IEEE*, Jyoti Prakash Mishra, *Member, IEEE,* and Amritesh Kumar, *Member, IEEE*

***Abstract*—The effect of reduction in the terminal voltage of a grid-tied inverter together with minimization in total harmonic distortion (THD) of terminal voltage and current, in the presence of dead-band in the switching remains always challenging for the researchers. With high switching frequency, this issue increases by many folds. The grid-tied inverter used for harmonic compensation acts as a shunt active power filter (SAPF). The implementation of the dead-band elimination technique is based on output current polarity detection, suffers from multiple zero-crossing due to the generation of high frequency switching harmonics during SAPF operation. This paper deals with a dead band elimination methodology together with harmonic compensation using SAPF under non-linear and unbalanced loading conditions. Here, SAPF uses a dual fundamental component extraction (DFCE) scheme with modified complex coefficient filter (MCCF) to extract fundamental components for reference current generation to implement an indirect current control PWM scheme. MCCF extracts fundamental components even in the presence of dc offset and unbalancing at the input. For incorporating dead-band elimination in control, exact polarity detection is done separately through SAPF output current estimation using sensed load current, and modified CCF computed reference source current. Since the estimated SAPF output current contains no high-frequency switching harmonics, the dead-band elimination technique can be correctly implemented. To synchronize the zero-crossing of the estimated and actual SAPF current, an adaptive delay control is used. The proposed MCCF based control logic's performance together with the dead-band elimination technique is validated for SAPF operating under non-linear, unbalance, and dynamic load conditions. The proposed dead-band elimination scheme based SAPF model is developed and simulated using MATLAB/ SIMULINK. The simulation results show a significant reduction in source current harmonics due to the elimination of odd harmonic injection into the system by incorporating dead-band elimination in SAPF operation. The real-time simulation of the proposed method is also validated using OPAL-RT OP4510 real-time platform.**

***Index Terms*—Shunt Active Power Filter (SAPF), Dead Band Elimination, Dual Fundamental Component Extraction (DFCE), Modified CCF (MCCF)**

I. Introduction

The proliferation of power electronics converters usage in the distribution system is the primary factor behind harmonic pollution. Harmonics in the distribution system create problems like interference, malfunctioning of sensitive devices, and heating. Harmonics can be reduced by using passive or active power filtering [1], [2]. Passive filtering methodology is frequency selective in nature. Therefore to compensate for varied harmonics components, an active power filtering technique is being used. This active filter is utilized for compensation of either voltage or current harmonics. Particularly, active filter housing voltage source converter (VSC) employing as current harmonic compensator is connected at the point of common coupling (PCC) in parallel with the load. The shunt connected VSC is responsible for compensating the harmonics in the system together with reactive power supply [3]. Diode rectifier with R-L load (current fed load) actually behaves as a source of harmonic current. For such current fed load, shunt connected active power filter is used to reduce the current harmonic level. Similarly, rectifier with parallel RC load is a voltage fed type of harmonics generator, which is a great source of current harmonics that may produce voltage distortion at PCC due to the presence of line impedances. To solve the issues related with a voltage fed type of harmonics, series active power filter or unified power quality conditioner (UPQC) is a better solution. On the other hand to provide limited voltage fed harmonics compensation with shunt active filter, appropriate value of AC or DC choke may be connected with diode rectifier [4]–[6].

To operate the VSC for proper compensation, the generation of proper switching signals is essential. For this, an appropriate reference current computation scheme using dominant time-domain-based approaches like synchronous reference frame (SRF) and instantaneous power (PQ) theory are used in literature [7]. Reference current generation based on instantaneous power theory (PQ) for SAPF control is simple, but control action suffers under corrupted PCC voltage due to switching harmonics, unbalance loading. Further, the use of low pass filter in the PQ control logic delayed the response. Moreover, in the SRF based control logic, PLL and LPF are used, which again affect the transient response. Others have also proposed an adaptive filter based control method for reference generation, like LMS (Least Mean Square), PNLMS (Proportionate Normalized LMS), and QLMS (Quaternion-Valued LMS), but suffers from steady-state active error and dynamic oscillation [8]. In [9], Adaline based control strategy is proposed for

Authors are with the Department of Electrical Engineering, National Institute of Technology Silchar, Assam, 788010, India e-mail: (alok_rs@ee.nits.ac.in).

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:23:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3104828, IEEE Transactions on Power Delivery

TABLE I
ABBREVIATIONS AND ACRONYM

PLL	Phase Locked Loop	SAPF	Shunt Active Power Filter	THD	Total Harmonic Distortion
STF	Self Tuning Filter	CCF	Complex Coefficient Filter	MCCF	Modified Complex Coefficient Filter
DFCE	Dual Fundamental Component Extraction	SOGI	Second Order Generalized Integrator	GSI	Generalized Sinusoidal Integrator
HCC	Hysteresis Current Control	FLL	Frequency Locked Loop	GI	Generalized Integrator
PWM	Pulse Width Modulation	LPF	Low Pass Filter	FFT	Fast Fourier Transform
PCC	Point of Common Coupling	V_{PCC}	Voltage at PCC	I_{abcL}/ I_{load}	Load Current
$I_{L1(fund.)}$	Fundamental load current	$V_{sa/sb/sc}$	Supply Voltage	V_{dc}, $V_{dc-link}$	Voltage across DC-Link
V^*_{dc}	Reference DC-Link Voltage	$I^*_{abcS(Ref.)}$	Source Current Reference	I_{abcS}/I_{Source}	Source Current
'S_{1-6}'	Gate signal	$C_{dc-link}$	Capacitance of DC link	L_C	Interfacing Inductance
R_f, C_f	Filter's Resistance and capacitance	R_S, L_S	Line Impedance Parameter source side	R_l	Resistive load at PCC
R_{Load}, L_{Load}	Load Impedance across rectifier	$I_{VSC}/I_{Comp.}$	Compensation current	G_α, G_β	α, β component of voltage /current input at CCF & MCCF
$G_{\alpha/\beta(fund.)}$	Fundamental output component of CCF/MCCF	G_{offset}	DC offset input at CCF/MCCF	k	Gain in filter
$G_{\alpha/\beta,-1(fund.)}$	Fundamental negative sequence component of CCF/MCCF	ω	Supply angular Frequency	$K^{'}$	Gain for offset loop
PSI	Pure Sinusoidal Integrator	I_{dc}	DC-link current	e_{dc}	Error in DC-link Voltage
k_p/k_i	Proportional/Integral gain of PI controller	$V_{o,error}$	Error in output of VSC due to dead band	T_{dead}	Dead-time
f_s	Switching frequency	t_{sett}	Settling time	T_d	Time delay
VSC	Voltage Source Converter	I^*_{VSC}	Estimated VSC current	R_{LL}, L_{LL}	AC choke parameter

SAPF, with reduced sensors.

Recently researchers have also advocated a DFCE with STF as an alternative to the conventional methods for the proper functioning of SAPF under ideal and abnormal grid voltage conditions [10], [11]. This approach has the benefit of being a faster response as it does not require PLL and LPF. Other than this, DFCE based reference current generation is not affected by distorted grid voltage. In [11], SRF theory is advocated in conjunction with DFCE for appropriate harmonic compensation. In control, both PLL and LPF exist similar to conventional SRF theory. So, it requires high computational effort. In [8], [10], STF based DFCE scheme is utilized for generation of synchronizing component with respect to balanced PCC voltage. However, under unbalanced PCC voltage, the method fails to generate synchronizing component. Further, when STF is applied to the sensed input voltage and current signal with DC bias, its performance gets affected. In [12], complex coefficient filter (CCF) based scheme is proposed to eliminate the DC biasing from input signal with fundamental component extraction in three phase system. CCF performance is satisfactory with a balanced supply, but with an unbalanced supply, performance deteriorates.

Incorporating the appropriate control scheme discussed above, VSC switches are operated with proper gating either using hysteresis current control or voltage control scheme. Either of the schemes generates PWM signals, which is ultimately used for switching of VSC [13]. After generating appropriate switching pulses, a dead band is provided between the top and bottom devices switching of each leg to avoid shoot-through [14]. However, due to dead band incorporation, power quality gets affected together with the decrease in the output of the fundamental component of VSC. Moreover, with dead-band control, some extra control logic, either through hardware or software, is required that further decreases the response time. With the usage of wide band-gap semiconductor switching devices in VSC for high switching operation, dead-band inclusion further affects the output voltage magnitude and power quality.

To nullify the effect of dead time, various dead-band compensation and elimination methods are suggested in literature [15]–[17]. In [18], a compensation scheme is incorporated, where the error is calculated between actual voltage and the reference voltage, then according to the error, the compensating voltage is added in the controller. In [19], error in voltage due to dead-time effect is calculated based on the charging and discharging of output capacitance of VSC. The results are saved in the look-up table and used as the compensating voltage in the controller. However, this dead time compensation scheme entirely depends upon VSC output current polarity detection. In [15], turn-off transition-based compensation method (TTCM) is utilized, where compensating voltage is calculated by turn off current and voltage error value. In [16], for dead-band elimination, grid current feedback and modulating function are used for the dead-band elimination requirement and use the current polarity detection method. In [17], VSC current polarity is determined using a diode conduction detector, digital signal processor, and complex programmable logic device (CPLD). The measurement of the currents in the bypass diodes of VSC is performed, and accordingly, PWM signals are generated to eliminate the dead time. In [20], the average value of VSC current is calculated around zero crossing point to detect VSC current's polarity.

In all dead time elimination schemes for VSC, switching pulses are generated for upper or lower switches according to VSC current's polarity or detection of diode current conduction. In all methods discussed above, complicated control is involved, and a few separate hardware is required, like DCD (diode current detection) to find the polarity of VSC current. In [21], a double second-order generalized integrator frequency locked loop (DSOGI-FLL) is used to detect output

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:23:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3104828, IEEE Transactions on Power Delivery

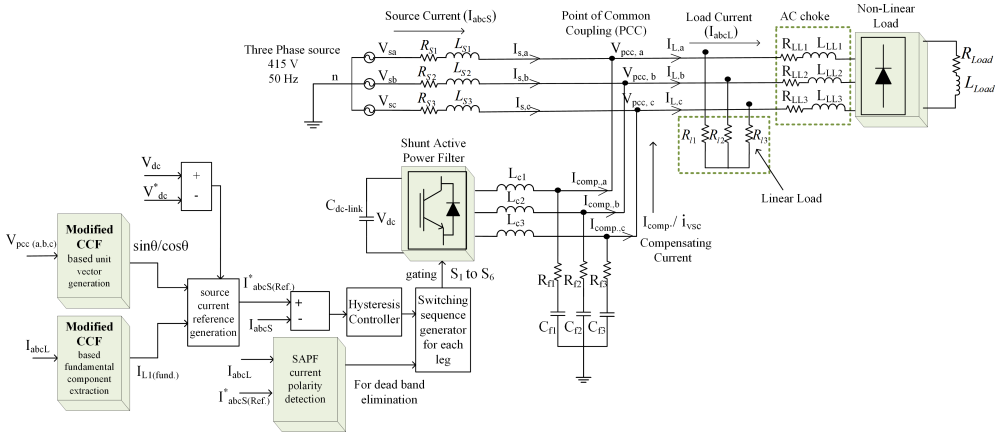


Fig. 1. Key elements of proposed MCCF based Shunt active power filter (SAPF) with dead time elimination technique

current polarity under distorted grid supply (PCC) and unbalanced loading conditions. These schemes do not apply to SAPF because SAPF compensates load harmonics by injecting harmonics to PCC [1], [22]. If compensating current is filtered out by DSOGI-FLL to apply dead-band elimination, the zero-crossing point is shifted, and compensation gets affected.

In the proposed work, a dead-band elimination technique embedded with MCCF based dual fundamental component extraction scheme is implemented in SAPF operation for harmonic and unbalance compensation. The MCCF is used to generate the synchronizing component with respect to PCC voltage. With the aid of another MCCF, fundamental component extraction from balanced/ unbalanced non-linear load current is also done. This dual fundamental component extraction scheme for reference current generation is less complex to implement. The major challenge in implementing dead-band elimination is the polarity detection near zero-crossing due to the presence of high-frequency switching noise in the SAPF compensation current. To incorporate dead band elimination in the proposed control, the exact polarity of SAPF output current is estimated separately in control logic using sensed load current, and MCCF computed reference source current. Because the estimated SAPF output current comprises no high-frequency switching harmonics, the dead-band elimination technique can be implemented effectively. To track the actual SAPF current's exact polarity, the estimated output current of SAPF and adaptive delay compensation are used. MCCF effectively removes the dc biasing and unbalancing in the input. Real-time dynamic loading performance is also up to the mark using MCCF. Under both testing conditions (unbalance and dynamic loading), the proposed dead-band elimination scheme gives the desired performance. To check the proposed scheme's effectiveness, the SAPF using an appropriate developed control scheme is simulated in MATLAB. Further, the proposed scheme's real-time simulation performance is also validated using the OPAL-RT OP4510 based real-time platform.

II. System Configuration

The proposed MCCF based SAPF and key elements are shown in Fig. 1. SAPF is connected to the grid at the point of common coupling (PCC) through interfacing inductor L_C. The function of the proposed strategy is to provide requisite compensation, with respect tothe load harmonics; switching harmonics (odd) through dead-band elimination; VAR requirement for UPF operation; load-side unbalance, and sensor-generated DC offset. To suppress voltage harmonics generated by VSC (two levels), a ripple filter tuned at switching frequency is connected as in Fig. 1. The voltage at DC-link capacitor (V_{dc}) of VSC is maintained at the reference value V^*_{dc} using SAPF with the proper control logic by taking requisite active power supply from the grid to compensate the losses in it. MCCF is connected for synchronization and filtration purposes while generating the reference source current. The computed three-phase balanced reference source current, along with the loss component of SAPF (to maintain the DC-link voltage constant), is passed through the hysteresis current controller (HCC) for the generation of required gating pulses (S_1 to S_6). At the load side, a combination of the non-linear and linear load is connected. Bridge type diode rectifier behaving as non-linear load. Also, there is unbalanced resistive load connected at PCC.

III. Control design

CCF is used to filter out the fundamental component from the input signals represented as 'G' in Fig. 2, where 'G' is voltage (V) or current (I) signals. The CCF is based on the generalized sinusoidal integrator (GSI). GSI is used to track the sinusoidal input. The primary control block diagram of GSI for single-phase system is shown in Fig. 2(b). Transfer function of the pure sinusoidal integrator (PSI) (enclosed by the dotted line in Fig. 2(b)) is expressed as (1) [23]:

$$\frac{G_{fund.}}{G_1} = \frac{1}{s - j\omega} \tag{1}$$

The complete transfer function of the GSI incorporating gain(k) is (2):

$$\frac{G_{fund.}}{G} = \frac{k}{s - j\omega + k} \tag{2}$$

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:23:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3104828, IEEE Transactions on Power Delivery

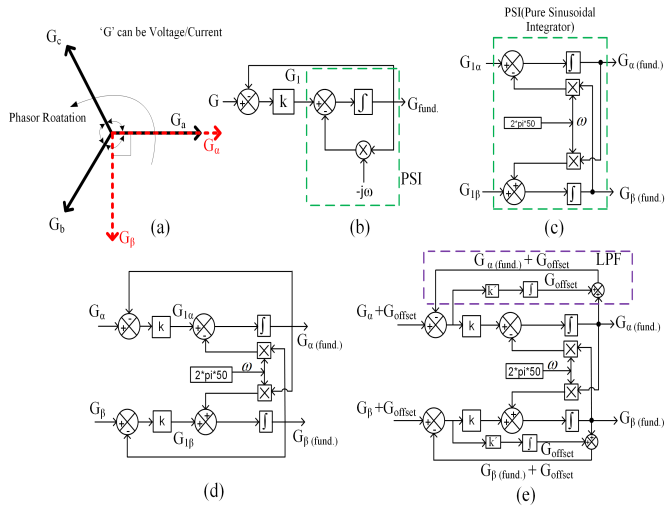


Fig. 2. (a) Phasor diagram of three phase input and corresponding α, β components, (b) Control block diagram of generalized sinusoidal integrator(GSI)(single phase), (c) Three-phase equivalent of pure sinusoidal integrator(dotted part of (b)), (d) Complex Co-efficient Filter(CCF) using concepts of GSI, (e) Improved CCF with DC offset rejection

The value of 'k' is constant, which is selected after tuning. 'ω' is the fixed angular frequency in rad/sec, equal to the input angular frequency. In GSI, due to the availability of only one signal in a single-phase system, a single loop is sufficient for fundamental extraction. However, in a three-phase system to apply GSI based logic, some modification is required. At first, the sensed three-phase signals are converted to orthogonal signals (G_α, G_β) using Clark's transformation matrix.

Here anti-clockwise phasor convention of three phase(a-b-c) phasor is considered and β is considered 90° lagging with respect to α component. The single-phase GSI control block is now generalized to a three-phase scheme, where the imaginary signal $(-j\omega G_{fund.})$ in Fig. 2(b) is replaced by coupling the orthogonal signal $(\omega G_{\beta(fund.)})$ in Fig. 2(c). Similarly, before feeding $G_{1\beta}$ to the integrator, $G_{\alpha(fund.)}$ is multiplied by 'ω' and then added to $G_{1\beta}$. Due to the availability of two integrator for individual α and β component, three-phase GSI is also called a double-integrator structure as in Fig. 2(c). The complex transfer function of the double-integrator structure as shown in dotted line in Fig. 2(c) is given by (3).

$$G_{\alpha(fund.)} + jG_{\beta(fund.)} = \frac{G_{1\alpha} + jG_{1\beta}}{s - j\omega} \tag{3}$$

From (3) it is clear that transfer function for Fig. 2(c) has also the term $(s - j\omega)$, similar to (1). So it also has the property of sinusoidal integrator, and it can be used in the three-phase system for fundamental component extraction after Clark's transformation. The close loop complex transfer function of CCF control structure in Fig. 2(d) is derived as (4):

$$G_{\alpha(fund.)} + jG_{\beta(fund.)} = k\frac{(G_\alpha + jG_\beta)}{s - j\omega + k} \tag{4a}$$

$$G_{\alpha(fund.)} + jG_{\beta(fund.)} = k\frac{(G_\alpha + jG_\beta)(s + k + j\omega)}{(s + k)^2 + \omega^2} \tag{4b}$$

The transfer function of the CCF is given by equation (4),

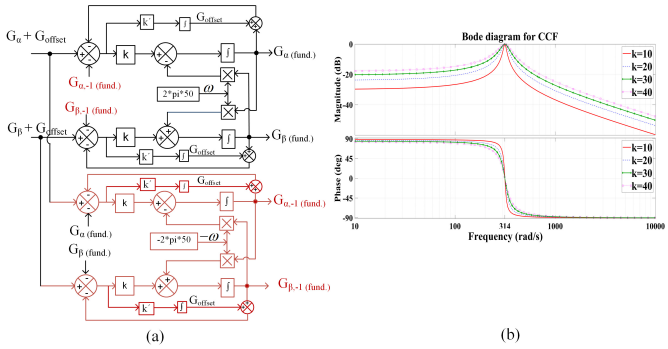


Fig. 3. (a) Control structure of MCCF with DC offset, unbalancing removal, (b) Bode plot depicting response under various gain values for MCCF

for which a bode plot is shown in Fig. 3(b). CCF has a high gain value at the fundamental frequency. In the Bode diagrams of Fig. 3(b), the influence of the gain 'k' on the frequency response can be seen. CCF has less attenuation for non-fundamental frequencies for high values of 'k'. Due to that, CCF is passing sub-harmonics and harmonics in the network for the high value of gain 'k'. Based on various response, the optimum value of 'k' is selected as 20.

A. Control operation of modified CCF for fundamental component extraction

1) Analysis of CCF under unbalanced input condition

To analyze the effect of unbalanced loading on the performance of CCF, a complex test input having different magnitude is taken as:

$$y(t) = y_\alpha(t) + jy_\beta(t) = \overbrace{Acos(\omega t + \phi)}^{G_\alpha} + \overbrace{jBsin(\omega t + \phi)}^{G_\beta} \tag{5}$$

where, y, A, B, ω and ϕ are the input signal, amplitude of α component, amplitude of β component, angular frequency, and initial phase of the input signal. Here 'A' and 'B' may be or may not be equal in amplitude.

Laplace transform of (5) gives,

$$Y(s) = L[Acos(\omega t + \phi) + jBsin(\omega t + \phi)] \tag{6}$$

After passing input signal Y(s) through CCF (Fig. 2(d)), which has the transfer function as (4) gives,

$$G(s)_{CCF} = L[Acos(\omega t + \phi) + jBsin(\omega t + \phi)]\frac{k}{s - j\omega + k} \tag{7}$$

Now the performance of $G(s)_{CCF}$ is studied under two conditions:

(i) Under equal magnitude of 'A' and 'B' (A = B), and by taking the inverse Laplace of (7) gives,

$$g(t)_{CCF} = A(1 - e^{-kt})cos(\omega t + \phi) + jA(1 - e^{-kt})sin(\omega t + \phi) \tag{8}$$

In the steady state, when $t \to \infty$, and simplifying (8) gives,

$$g(t)_{CCF} = Acos(\omega t + \phi) + jAsin(\omega t + \phi) = y(t) \tag{9}$$

From (9) it is clear that, CCF output is same as input signal in the steady state conditions. Means, CCF is operating as

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:23:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3104828, IEEE Transactions on Power Delivery

amplitude integration of the input signal and tracks the input signals of fundamental angular frequency 'ω'. Moreover, it is clear that Fig. 2(c) may be used with complex input signal for CCF.

(ii) Performance of CCF under unequal amplitude (A ≠ B), by inverse Laplace of (7) gives,

$$g(t)_{CCF} = L^{-1}[L[Acos(\omega t+\phi)+jBsin(\omega t+\phi)]\frac{k}{s-j\omega+k}] \tag{10}$$

$$\begin{aligned} g(t)_{CCF} = {} & L^{-1}[L[Acos(\omega t+\phi)]\frac{k}{s-j\omega+k}]+ \\ & +L^{-1}[L[jBsin(\omega t+\phi)]\frac{k}{s-j\omega+k}] \end{aligned} \tag{11}$$

Apply convolution theorem for Inverse Laplace,

$$\begin{aligned} g(t)_{CCF} = {} & Ak\left(\int_0^t cos(\omega u+\phi)e^{(j\omega-k)(t-u)}\,du\right)+ \\ & +jBk\left(\int_0^t sin(\omega u+\phi)e^{(j\omega-k)(t-u)}\,du\right) \end{aligned} \tag{12}$$

After simplification of (12) gives,

$$\begin{aligned} g(t)_{CCF} = {} & \frac{k}{(k-j\omega)^2+\omega^2}[\omega sin(\omega t+\phi)(A+B) \\ & +Akcos(\omega t+\phi)-e^{-(k-j\omega)(t)}(Akcos\phi+A\omega sin\phi+B\omega sin\phi)] \\ & +\frac{jk}{(k-j\omega)^2+\omega^2}[-\omega cos(\omega t+\phi)(A+B) \\ & +Bksin(\omega t+\phi)-e^{-(k-j\omega)(t)}(Bksin\phi-B\omega cos\phi-A\omega cos\phi)] \end{aligned} \tag{13}$$

In the steady state, when $t \to \infty$, (13) can be written as,

$$\begin{aligned} g(t)_{CCF} = {} & \overbrace{\frac{k}{(k-j\omega)^2+\omega^2}}^{\mathrm{R}}\overbrace{[\omega sin(\omega t+\phi)(A+B)}^{\mathrm{P}} \\ & +Akcos(\omega t+\phi)]+ \\ & +\frac{jk}{(k-j\omega)^2+\omega^2}\overbrace{[-\omega cos(\omega t+\phi)(A+B)}^{\mathrm{Q}}+ \\ & +Bksin(\omega t+\phi)] \end{aligned} \tag{14}$$

$$g(t)_{CCF} = R\overbrace{[P+Akcos(\omega t+\phi)]}^{g(t)_\alpha}+jR\overbrace{[Q+Bksin(\omega t+\phi)]}^{g(t)_\beta} \tag{15}$$

Where, R = $\frac{k}{(k-j\omega)^2+\omega^2}$, P = $\omega sin(\omega t+\phi)(A+B)$, Q = $-\omega cos(\omega t+\phi)(A+B)$.

From (15) it is clear that $g(t)_\alpha$ component contains extra term 'P' and $g(t)_\beta$ component contains extra term 'Q', apart from fundamental component in the steady-state. So, CCF could not filter out the fundamental balanced component under the unbalanced condition of input signals.

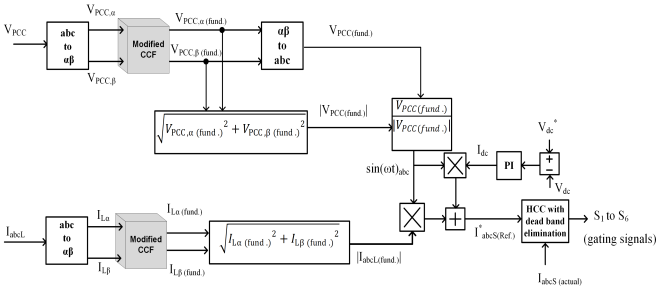


Fig. 4. MCCF based dual fundamental component extraction methodology for harmonic compensation and dead band elimination

2) MCCF under unbalancing condition

In the case of unbalanced loading, the sensed load current also becomes unbalanced. Reference source current generation using the conventional CCF through the sensed unbalanced load current ultimately produces unbalancing in the reference current. So, to avoid the unbalanced reference current generation, the negative sequence component in it needs to be removed. For this purpose, the fundamental negative sequence component is subtracted from sensed input signals as in Fig. 3(a). The fundamental negative sequence component is computed with another CCF tuned at negative sequence fundamental frequency (-ω) parallel with the positive sequence fundamental frequency tuned CCF. Further, to remove the dc offset from the sensed input signal, an additional loop containing generalized integrator tuned at zero frequency is used. The obtained value (G_{offset}) is again removed from the sensed signal. The complete block containing different control loops for MCCF is shown in Fig. 3(a).

3) MCCF based control logic for SAPF operation

Various signals such as PCC voltage, load current, source current, and dc-link voltage must be sensed in order to produce correct reference signals for HCC based SAPF as shown in Fig. 4. The sensed PCC voltage is fed to the MCCF module to obtain the fundamental component ($V_{PCC,\alpha(fund.)}, V_{PCC,\beta(fund.)}$), which are then used to generate the unit vector sin(ωt), called as synchronizing component. MCCF is also used to obtain the fundamental load current components ($I_{L\alpha(fund)}$, $I_{L\beta(fund)}$) from the sensed load currents. After a proper mathematical operation, these two components of load current evaluate the magnitude of the fundamental load current ($|I_{abcL(fund)}|$). To compute the active component of the load current needed to be supplied by the grid, the magnitude of the fundamental load current component ($|I_{abcL(fund)}|$) is multiplied by the synchronizing component. The reference source current ($I^*_{abcS(Ref.)}$) is mainly composed of this fundamental active load current component and an active loss component to overcome the losses occurring in SAPF. The loss is computed using (16), (17). Here, I_{dc} is the necessary dc-link current component to maintain the dc-link voltage at its reference value and compensate for SAPF losses.

The error in actual and reference dc-link voltage is computed as,

$$e_{dc}(k) = V_{dc}(k)^* - V_{dc}(k) \tag{16}$$

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:23:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3104828, IEEE Transactions on Power Delivery

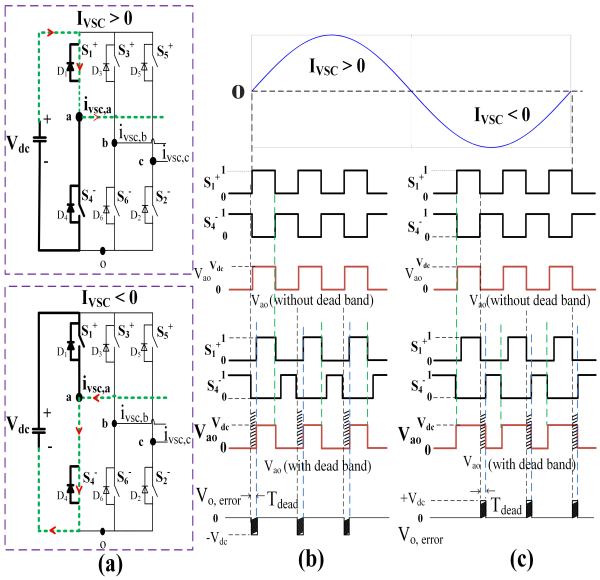


Fig. 5. (a) VSC operation for one leg, (b)Voltage error in output of VSC due to dead band effect in one leg, when $I_{VSC} > 0$ (c) when $I_{VSC} < 0$

The error is processed through digital PI controller to evaluate the required dc-link current (loss component) $I_{dc}(k)$ as:

$$I_{dc}(k) = I_{dc}(k-1) + k_i[e_{dc}(k) - e_{dc}(k-1)] + k_p[e_{dc}(k)] \quad (17)$$

B. Effect of dead band in VSC

Dead-band is created between top and bottom switching of same leg devices of VSC to avoid shoot through. Provision of such dead band in the switching control logic results in the lower output voltage of VSC as compared to the output voltage of VSC without dead band. The difference in VSC output voltage with and without dead-band gives voltage error along with lower order odd harmonics. Effect of incorporation of dead-band is shown in Fig. 5. Upper switches (S_1^+, S_3^+, S_5^+) are treated as positive group and lower areas negative group (S_4^-, S_6^-, S_2^-). To avoid the shoot-through, a dead time "T_{dead}" is given during switching between the positive group and negative group devices of the same leg. It is observed that due to the incorporation of turn-on delay, there is a reduction of output voltage during the positive half of VSC current and gain of output voltage during the negative half of VSC current. Therefore, the incorporation of dead time results in voltage error, creating current distortion and increasing the harmonics. These error voltages have odd symmetry with respect to the polarity of converter current and hence only odd harmonics, primarily the dominant $5^{th}, 7^{th}, 11^{th}$ order exist. The value of error voltage is derived as (18).

$$V_{o,error}(avg.) = V_{dc}T_{dead}f_s \quad (18)$$

where "f_s" is switching frequency of device, "T_{dead}" is dead time, "V_{dc}" is assumed DC-link voltage.

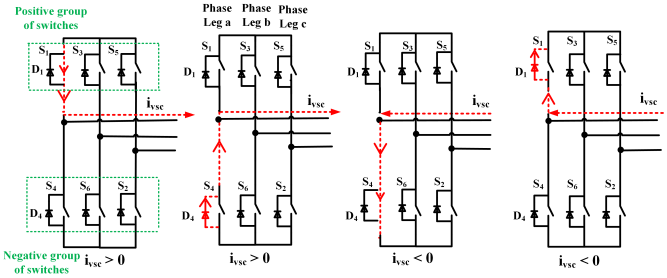


Fig. 6. Switching operation of three-phase inverter with dead-band elimination scheme

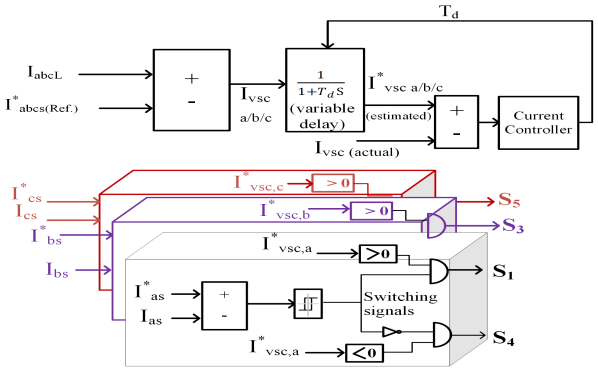


Fig. 7. Control strategy for SAPF current estimator and HCC operation with dead band elimination

C. Control logic for Dead-time elimination in VSC

With the VSC switching in the kHz range (around 20 kHz), there is a requirement of dead time of 1-5 μsec between switching of positive and negative devices of each leg. However, this dead time creates more harmonics in the system, and extra hardware is required to provide the dead time. From Fig. 5 (b), it is clear that with positive VSC current (for phase 'a'), either flows through switch 'S_1^+' when it is enabled or through diode D_4 when 'S_1^+' is OFF. It means that with a positive current, there is no requirement of switching 'S_4^-' at all. Similar concepts are applicable with the negative direction of VSC current. Hence there is no dead-band requirement between positive and negative switches if the switch is operated according to VSC current polarity. Therefore, the root of dead time elimination methods lies in the switching of either upper or lower devices of one leg according to the polarity of VSC output current as given in (19), (20) and uses other leg devices for return path as shown in Fig. 6.

$$Positive\ Switch = Switching\ signal * sign(i_{vsc}) \quad (19)$$

$$Negative\ Switch = \overline{Switching\ signal} * \overline{sign(i_{vsc})} \quad (20)$$

According to (19) and (20), positive and negative switches operation is controlled by the polarity of VSC output current. Throughout the half of time period when $sign(i_{vsc}) > 0$, the negative group of switch is in the non-operating condition. Similar is the case for positive group of switches when $sign(i_{vsc}) < 0$.

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:23:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3104828, IEEE Transactions on Power Delivery

TABLE II
PARAMETER VALUES IN MCCF BASED SAPF

Mains supply	415(Ph-Ph) V & 50 Hz
DC link voltage(V_{dc})	700 V
DC link capacitor($C_{dc-link}$)	1500 μF
Interface inductor (L_c)	3 mH
Ripple filter(R_f and C_f)	200 Ω, 4 μF
Rectified fed load(R_{Load} and L_{Load})	20 Ω, 15 mH
Line impedance (R_S and L_S)	0.05 Ω, 0.05 mH
AC Choke (load side)(R_{LL} and L_{LL})	0.2 Ω, 1.5 mH
Unbalance load (R_{l1}, R_{l2}, R_{l3})	30 Ω, 20 Ω, 10 Ω

Therefore, the reliability of the dead-band elimination scheme lies in the accurate detection of VSC current polarity. However, with SAPF, this challenge of polarity detection increases many folds. SAPF injects the harmonics in the anti-phase of load harmonics to make the source current harmonic free. Other than this, SAPF current is also polluted by high frequency switching noises of devices. This causes multiple zero crossings in the SAPF current, making it difficult to accurately detect its polarity. To avoid this, the actual SAPF current is not directly measured; instead, it is estimated. The phase difference error between the estimated and actual VSC current is minimized quickly through closed-loop delay-time constant computation in the VSC current estimator to achieve the desired operation of VSC under dead-band elimination mode.

From the total demanded load current (I_{Load}), $I^*_{abcS(Ref.)}$ component is supplied from grid and rest is supplied from VSC as,

$$i_{vsc} = I_{abcL} - I^*_{abcS(Ref.)} \tag{21}$$

$$\text{where, } I^*_{abcS(Ref.)} = (I_{dc} + |I_{abcL(fund.)}|)sin(\omega t)_{abc} \tag{22}$$

Fig. 7 shows the control strategy, which estimates the SAPF current to avoid switching harmonics caused by high frequency switching of its devices. DFCE control technique compute the source current reference though MCCFs. From the total load current demand, other than this reference source current, is supplied through VSC only. The estimated VSC current leads the actual SAPF current because it is estimated directly without any propagation delay. To overcome this, optimized delay is provided during the process VSC current estimation, before gating signals generation using HCC. The estimated VSC current is used to operates the switching devices according to its polarity detection near zero crossing, to implement dead-band elimination as shown in Fig. 6. This scheme has the advantage of avoiding high frequency switching harmonics in the estimated SAPF current to detect its polarity accurately near the zero crossings, during implementation of dead time elimination technique.

IV. SIMULATION RESULTS

The proposed three-phase SAPF operation, with and without dead-band elimination, based on CCF and MCCF control strategies, while feeding non-linear (three phase diode bridge rectifier) and linear loads, is simulated using MATLAB Simulink environment. In both(with/ without dead band) cases, the harmonic analysis of source current is also performed. Table II shows the values of the various parameters of the proposed system. The PCC voltage, load current, source current, compensating current, and dc-link voltage are shown in Fig. 8 (a)-(e). Further, the performance of SAPF operations with and without dead band elimination are compared as shown in Table III. The proposed strategy of dead-band elimination with MCCF is also tested with unbalance loading at PCC, created through resistances in phase-a,b,c with values 30 Ω, 20 Ω, 10 Ω respectively. Similarly, to test the dynamic loading performance, the rectifier fed load is provided with step-change by connecting 20 Ω in parallel combination with existing 20 Ω, 15 mH (R_{Load} and L_{Load}) load.

A. Performance of SAPF with dead-band elimination:

In Fig. 8 (a)-(e) shows the SAPF operation using dead-band elimination with PCC voltage, load current, source current, compensation current, dc-link voltage respectively. Initially, the DC-link capacitor of SAPF is charged up to 700 V to avoid the transient during the turning ON process of SAPF. Three-phase grid supply voltage (V_{PCC}) is balanced-pure sinusoidal, as in Fig. 8 (a). As the loading is unbalanced in nature, load current (I_{load}) is unbalanced, as in Fig. 8 (b). Due to the effect of non-linearity and unbalancing in load, source current is polluted with harmonics and unbalance as shown in Fig. 8 (c) (up to 0.05 s) with peaks of 41 A, 46 A, and 47 A. At t = 0.05 s, SAPF operation is enabled with only CCF-based control logic. As compensation is started at t = 0.05 s, hysteresis current controller-based SAPF compensates the harmonic currents to make the source current sinusoidal at unity power factor. To apply dead-band elimination in VSC, gating signals are applied between upper and lower switches according to the polarity of the estimated VSC current. When SAPF starts compensating the harmonic demand of load at t = 0.05 s, the source current harmonics reduces. However, still source current possesses unbalancing up to t = 0.25 s, as CCF-based control logic could not extract the fundamental balanced component from unbalanced load current during reference source current generation.

After t = 0.25 s, MCCF is utilized to extract a balanced fundamental load current component, which is further used for reference source current generation. Therefore, the use of the proposed MCCF helps to maintain source current (peak ≈ 45 A) and sinusoidal in the presence of non-linear and unbalanced load. The stiffness of the proposed compensator using MCCF is also validated under step change in the load at t = 0.35 s to test its proper functioning under a dynamic loading environment. With the change in loading, SAPF successfully mitigates the harmonics and quickly reaches the steady-state condition.

B. Comparison of SAPF with and without dead-band elimination

Source current (I_{Source}) THD before compensation is 15.07% as shown in Fig. 10, which is higher than the as per the IEEE-519 standard (5%) for THD in grid supply current. In the proposed circuit, ratio of short circuit current and load current is approximately equal to 13 (< 20), therefore according to IEEE-519-2014 standard, requirement of THD is

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:23:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3104828, IEEE Transactions on Power Delivery

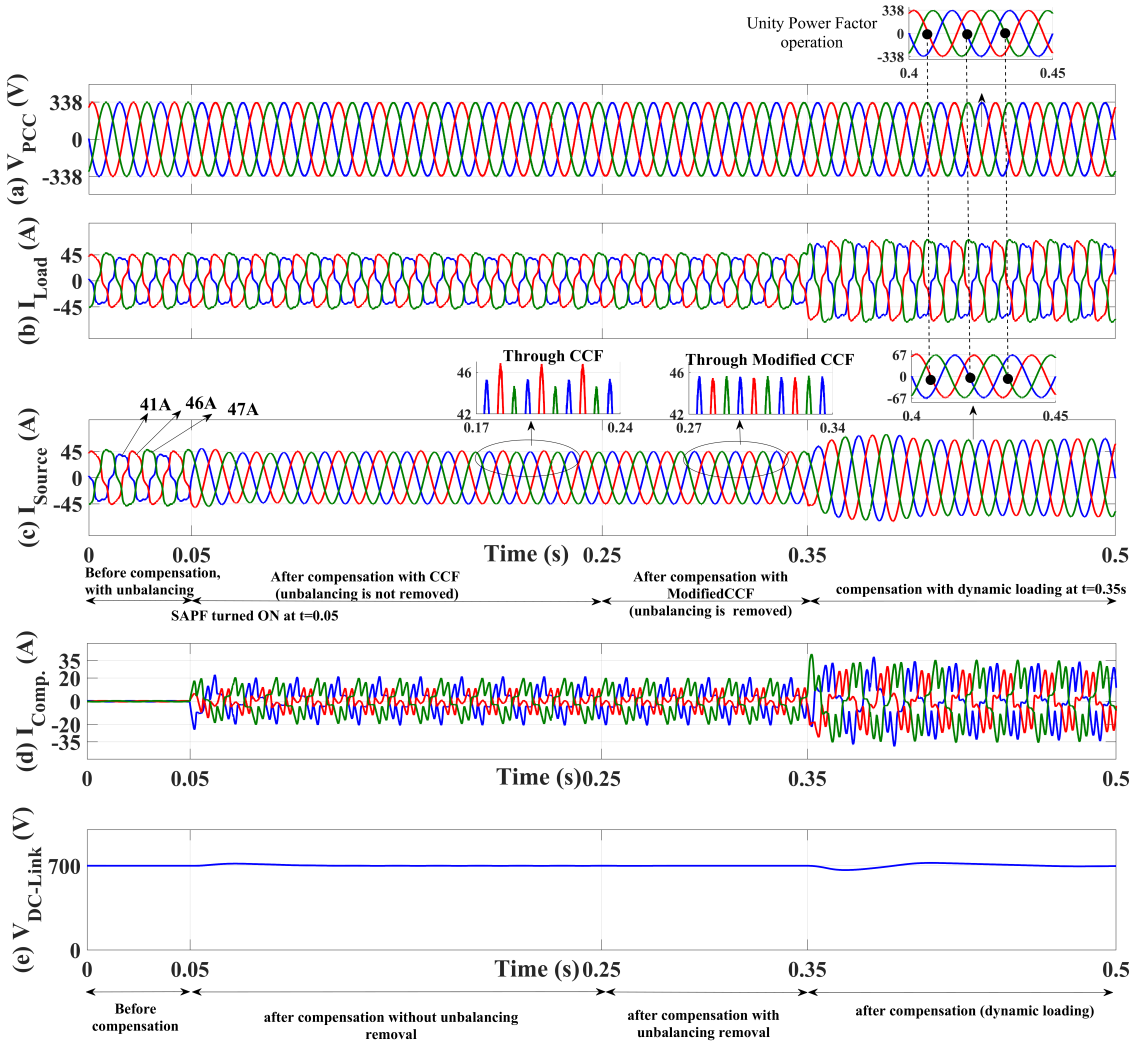


Fig. 8. With dead-band elimination-SAPF, the simulation plot for: (a) point of common coupling voltage, (b) load current, (c) source current, (d) compensation current, (e) dc-link voltage

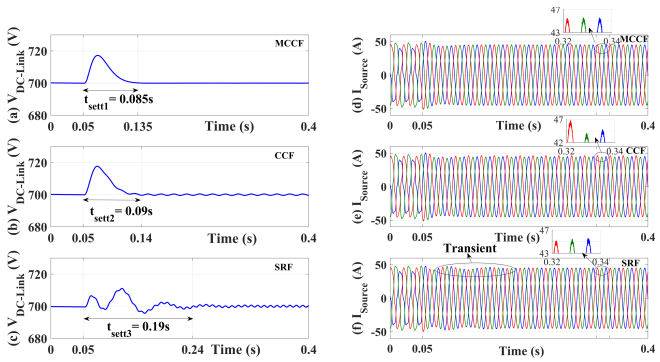


Fig. 9. (i) Dynamic response of dc-link voltage in SAPF using: (a) MCCF, (b) CCF, (c) SRF, (ii) Response of source current in SAPF using: (d) MCCF, (e) CCF, (f) SRF

below 5%. With the dead-band elimination scheme, %THD is 0.59%, which is within the IEEE-519 standard. The proposed dead-band elimination based HCC switching strategy uses a separate SAPF current estimator to avoid the effect of multiple zero crossings in the presence of high frequency switching harmonics in the actual SAPF current.

When the SAPF is operated with a dead-band of 5 μs, the measured source current THD is 2.29% after compensation as shown in Fig. 11, which is nearly four times higher than that with the dead time elimination technique. Although the source current harmonics are within the range of the IEEE-519 standard using the dead-band based HCC operated SAPF, it has some following demerits:

- As the dead time increases, the source current THD level deteriorates more.
- Extra circuit is required to take care of the dead-band.
- With the use of high switching frequency devices, the SAPF output voltage error also increases.
- With the high value of SAPF input DC voltage, the effect of dead-band becomes more dominating and increases THD level of source current.

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:23:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3104828, IEEE Transactions on Power Delivery

TABLE III
PERFORMANCE COMPARISON OF SAPF WITH,(A) DEAD-BAND ELIMINATION (B) WITH DEAD-BAND

Technique used	% THD before compensation	% THD after compensation	Error in output voltage	Suitability to high switching frequency	Extra Hardware requirement for dead-band
(a) With dead-band elimination	15.07	0.59	NO	YES	NO
(b) With dead-band	15.07	2.29	YES	NO	YES

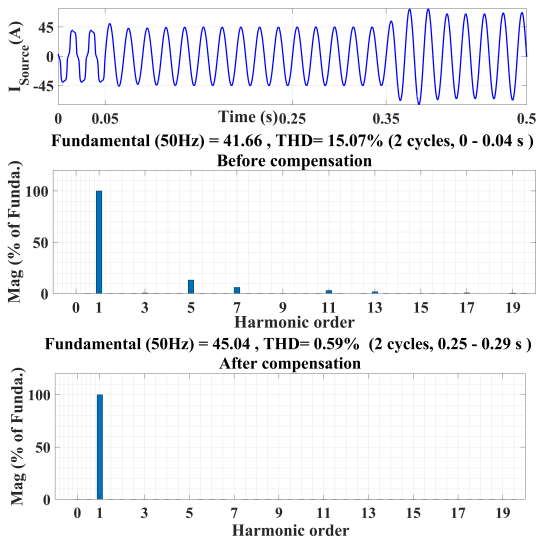


Fig. 10. THD analysis of MCCF based SAPF with dead-band elimination based HCC PWM strategy

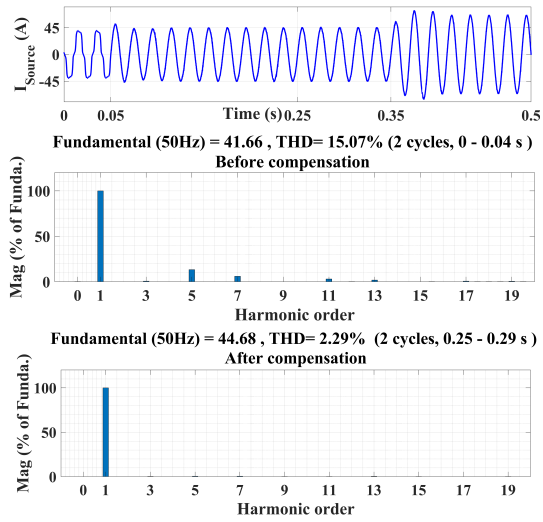


Fig. 11. THD analysis of MCCF based SAPF without dead-band elimination based HCC strategy

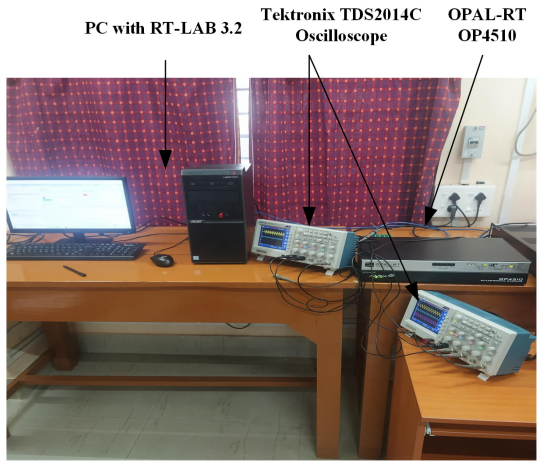


Fig. 12. Setup for real-time simulation to verify the accuracy of MCCF-based SAPF with dead-band elimination

C. Performance comparison of SAPF with MCCF, CCF, and SRF based reference source current generation

To identify the proposed control scheme's stiffness, DC-link voltage across VSC and source current are captured for SAPF using different control logic like MCCF, CCF, and SRF, in simulation as shown in Fig. 9. It is visible using SRF based SAPF operation, dc-link voltage response is more oscillatory with high settling time (0.19 s). In the present case, SRF based control scheme with 8th ordered Butterworth LPF is used for filtering purposes. Performance of SAPF with SRF based control varies with the type, order of LPF, and PLL. With the CCF-based SAPF operation, the dc-link voltage response is settling time is reduced to 0.09 s, but it has a ripple component because unbalancing is not removed as observed in the source current signal. MCCF based SAPF operation is a suitable choice, as it shows the lowest settling time (0.085 s) with minimum ripple in the dc-link voltage response and successfully removes the unbalancing in source current.

V. REAL-TIME SIMULATION RESULTS

The performance of dead-band elimination based SAPF operation with MCCF is evaluated through real-time simulation using OP4510 platform are shown in Fig.13-15. The Real-time simulation results are captured on Tektronix TDS2014C, 100 MHz, 2-GS/s oscilloscope through analog out-port of the real-time simulator as shown in Fig. 12. In Fig. 13, the performances of the proposed system are validated under step load change (dynamic load) at t = 0.125 s. In Fig. 13, the proposed system performance is also validated under unbalanced loading conditions (case II).

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:23:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3104828, IEEE Transactions on Power Delivery

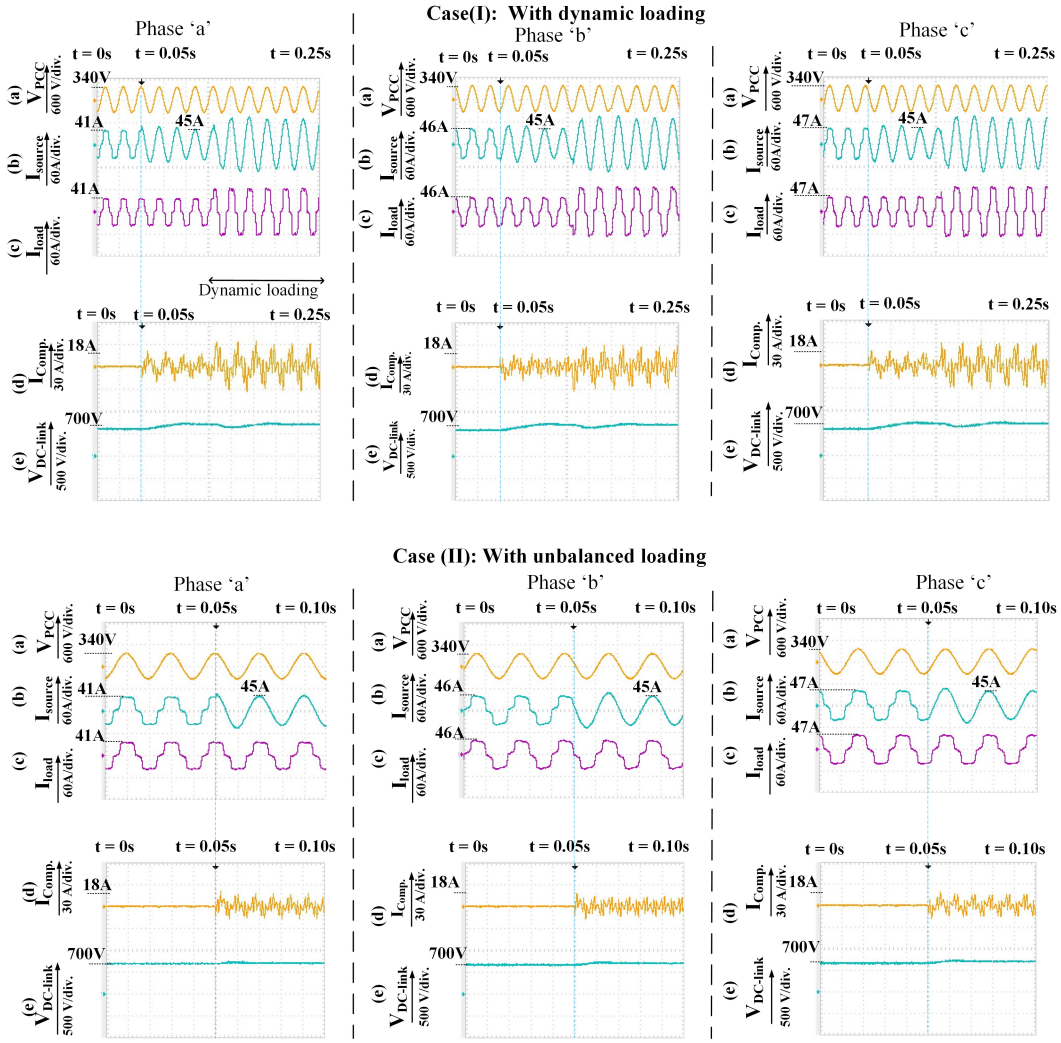


Fig. 13. Real-time simulation results with dead-band elimination based SAPF for dynamic loading(case I) and unbalanced loading(case II) condition: (a) point of common coupling voltage(V_{pcc}- 600V/div.), (b) source current (I_{source}- 60A/div.), (c) load current (I_{load}- 60A/div.), (d) compensation current ($I_{comp.}$- 30A/div.), (e) dc-link voltage ($V_{DC-Link}$- 500V/div.)

A. *Real-time simulation of SAPF with dead-band elimination under non-linear and dynamic loading condition*

In this case, the source current compensation is validated under dynamic loading condition through real-time simulation of proposed scheme. With the triggering pulses are applied to SAPF, the compensated source current is observed in the scope, as shown in Fig. 13(b) (case I) for each phase. It is observed that the real-time simulated results resemble that of MATLAB simulation results. In Fig. 13(c), the load current is captured, which is highly non-linear due to rectifier load. In Fig. 13, from t = 0 s to t = 0.05 s the SAPF is not triggered, which results in non-sinusoidal source current. After t = 0.05 s, source current is sinusoidal through compensation current as shown in Fig. 13(b). DC-link voltage is almost settled at a reference voltage of 700 V as shown in Fig. 13(e). With the step change in load, source and load current magnitude increases without affecting the compensation characteristics of the proposed SAPF. For a better understanding of dynamic loading performance, all the three-phase source currents are captured simultaneously as shown in Fig.14(a). The peak of the source current is increased to 67 A with the dynamic loading.

B. *Real-time simulation of SAPF with dead-band elimination under non-linear unbalanced loading condition*

To test the unbalancing rejection capability of MCCF, the performance of the proposed system is validated under an unbalanced loading condition. Source current is non-linear and unbalanced before the triggering pulses are applied to SAPF at t = 0.05 s, as shown in Fig. 13 (case II). After t = 0.05 s, the source current becomes almost balanced and harmonics free, as shown in Fig. 13(b). Fig. 13(d) shows the injected compensation current before and after the activation of SAPF operation. DC-link voltage maintained at reference value as shown in Fig. 13(e). For more clarity of unbalancing removal by MCCF, source current is captured in Fig.14(b) with all three-phase simultaneously. Peak values of each phase in source current are 41 A, 46 A, and 47 A before starting

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:23:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3104828, IEEE Transactions on Power Delivery

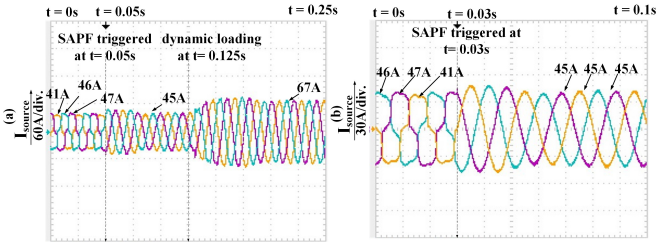


Fig. 14. Source current (I_{source}) in MCCF based SAPF before and after compensation with : (a) dynamic loading condition (b) unbalanced loading condition

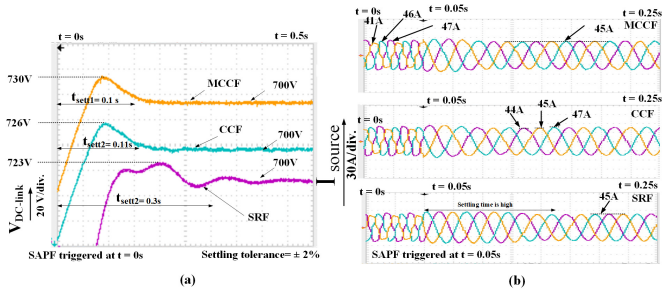


Fig. 15. (a) Dynamic response of $V_{DC-Link}$- (20V/div.) (b) Source current (I_{source}), with MCCF, CCF and SRF based control strategy for SAPF reference signal generation in real -time simulation

of compensation. After switching ON the compensation using MCCF, unbalancing is removed, and the source current peak is maintained at 45 A.

C. Performance comparison of SAPF with MCCF, CCF, and SRF based reference current generation in real-time simulation

Transient response of SAPF is based on settling of DC-link voltage of VSC and characteristics of its source current. In Fig.15 (a) and Fig. 15 (b), the transient responses of DC-link voltage are shown with three schemes (MCCF, CCF, and SRF) of reference source current generation along with their corresponding source current responses. To validate the transient response, all three schemes are tested simultaneously with the same parameter. All three SAPFs are triggered at t = 0 s, and corresponding settling time is recorded (with a tolerance band of ±2%). The settling time of dc-link voltage with MCCF based SAPF is 0.1 s, and DC-link voltage has smooth, steady curves after settling point. The DC-link voltage of SAPF using CCF-based control has a settling time of 0.11 s but shows the ripples in it even after the settling point is reached. The DC-link voltage of SAPF with SRF based control has more oscillations, which is settled at 0.3 s. Also, it has high transient with the variation of LPF tuning and PLL.

VI. Conclusion

In this paper, the proposed dead-band elimination based SAPF operation with DFCE method embedded with MCCF, for source current harmonic compensation, is found compelling at the high switching frequency. MCCF structure removes unbalancing smoothly and is fast enough to cope up with dynamic loading. The proposed method is less complex as compared to other schemes of compensation. As the estimated VSC current of the SAPF does not contain any high frequency switching harmonics, it accurately detects its polarity near zero crossings and more straightforward implementation of the dead-band elimination technique for SAPF. There is an improvement in the THD level of source current is also observed with dead time elimination. Theoretically, the proposed switching scheme is adaptive and robust to any mismatch of switching device characteristics. The scheme also eliminates the voltage error due to the dead-band switching between the same leg devices of SAPF. The real-time simulation results resemble the MATLAB simulation results, hence verifying the proposed control strategy's stiffness.

Acknowledgments

Authors expressed their gratitude for the provision of OPAL-RT OP4510 based real-time digital simulator platform by National Institute of Technology Silchar and also thankful to Science and Engineering Research Board sponsored IMPRINT-IIC project IMP/2019/000234 for additional technical support.

References

[1] R. Kumar and H. O. Bansal, "Shunt active power filter: Current status of control techniques and its integration to renewable energy sources," *Sustainable Cities and Society*, vol. 42, pp. 574–592, 2018.

[2] L. Tey, P. So, and Y. Chu, "Improvement of power quality using adaptive shunt active filter," *IEEE transactions on power delivery*, vol. 20, no. 2, pp. 1558–1568, 2005.

[3] S. Biricik, S. Redif, Ö. C. Özerdem, S. K. Khadem, and M. Basu, "Real-time control of shunt active power filter under distorted grid voltage and unbalanced load condition using self-tuning filter," *IET Power Electronics*, vol. 7, no. 7, pp. 1895–1905, 2014.

[4] R. Pandurangan, P. Kaliannan, and P. Shanmugam, "Effects of current distortion on dc link inductor and capacitor lifetime in variable frequency drive connected to grid with active harmonic filter," *IEEE Transactions on Industry Applications*, vol. 57, no. 1, pp. 492–505, 2021.

[5] F. Z. Peng, "Application issues of active power filters," *IEEE Industry applications magazine*, vol. 4, no. 5, pp. 21–30, 1998.

[6] S. Ouchen, M. Benbouzid, F. Blaabjerg, A. Betka, and H. Steinhart, "Direct power control of shunt active power filter using space vector modulation based on supertwisting sliding mode control," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 9, no. 3, pp. 3243–3253, 2021.

[7] H. Akagi, "New trends in active filters for power conditioning," *IEEE Transactions on Industry Applications*, vol. 32, no. 6, pp. 1312–1322, 1996.

[8] B. Singh *et al.*, "Dfce based control of solar pv-wind energy system interfaced to distribution network," in *2018 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES)*. IEEE, 2018, pp. 1–6.

[9] S. Sharma, V. Verma, and R. K. Behera, "Real-time implementation of shunt active power filter with reduced sensors," *IEEE Transactions on Industry Applications*, vol. 56, no. 2, pp. 1850–1861, 2019.

[10] Y. Hoon, M. A. M. Radzi, M. K. Hassan, and N. F. Mailah, "Operation of three-level inverter-based shunt active power filter under nonideal grid voltage conditions with dual fundamental component extraction," *IEEE Transactions on Power Electronics*, vol. 33, no. 9, pp. 7558–7570, 2017.

[11] T. Ngo, S. Biricik, and M. Basu, "A self-tuning grid synchronization method for active power filters," *Electric Power Components and Systems*, vol. 44, no. 17, pp. 1947–1957, 2016.

[12] S. Golestan, J. M. Guerrero, J. C. Vasquez, A. M. Abusorrah, and Y. Al-Turki, "A study on three-phase flls," *IEEE Transactions on Power Electronics*, vol. 34, no. 1, pp. 213–224, 2018.

[13] L. Zhang, B. Gu, J. Dominic, B. Chen, C. Zheng, and J.-S. Lai, "A dead-time compensation method for parabolic current control with improved current tracking and enhanced stability range," *IEEE Transactions on Power Electronics*, vol. 30, no. 7, pp. 3892–3902, 2014.

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:23:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3104828, IEEE Transactions on Power Delivery

[14] C. Song and Z. Zeng, "Analysis and elimination of circulating current caused by dead time for a grid-connected inverter parallel system," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 13, no. 10, pp. 1520–1527, 2018.
[15] T. Mannen and H. Fujita, "Dead-time compensation method based on current ripple estimation," *IEEE Transactions on Power Electronics*, vol. 30, no. 7, pp. 4016–4024, 2014.
[16] Y. Wang, Q. Gao, and X. Cai, "Mixed pwm for dead-time elimination and compensation in a grid-tied inverter," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 10, pp. 4797–4803, 2011.
[17] L. Chen and F. Z. Peng, "Dead-time elimination for voltage source inverters," *IEEE transactions on power electronics*, vol. 23, no. 2, pp. 574–580, 2008.
[18] Y. Ji, Y. Yang, J. Zhou, H. Ding, X. Guo, and S. Padmanaban, "Control strategies of mitigating dead-time effect on power converters: An overview," *Electronics*, vol. 8, no. 2, p. 196, 2019.
[19] J. Schellekens, R. Bierbooms, and J. Duarte, "Dead-time compensation for pwm amplifiers using simple feed-forward techniques," in *The XIX International Conference on Electrical Machines-ICEM 2010*. IEEE, 2010, pp. 1–6.
[20] Y.-K. Lin and Y.-S. Lai, "Dead-time elimination of pwm-controlled inverter/converter without separate power sources for current polarity detection circuit," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 6, pp. 2121–2127, 2009.
[21] Q. Yan, R. Zhao, X. Yuan, W. Ma, and J. He, "A dsogi-fll-based dead-time elimination pwm for three-phase power converters," *IEEE Transactions on Power Electronics*, vol. 34, no. 3, pp. 2805–2818, 2018.
[22] N. Beniwal, I. Hussain, and B. Singh, "Control and operation of a solar pv-battery-grid-tied system in fixed and variable power mode," *IET Generation, Transmission & Distribution*, vol. 12, no. 11, pp. 2633–2641, 2018.
[23] Z. Xin, R. Zhao, P. Mattavelli, P. C. Loh, and F. Blaabjerg, "Re-investigation of generalized integrator based filters from a first-order-system perspective," *Ieee Access*, vol. 4, pp. 7131–7144, 2016.



Mr. Alok Kumar Dubey (S'20) completed B.E. in Electronics and Telecommunication Engineering from Chhatrapati Shivaji Institute of Technology, Durg, Chhattisgarh, India in 2010. He obtained M.E. in Power Electronics from Rungta College of Engineering and Technology, Bhilai, Chhattisgarh, India in 2013. From 2013 to mid June 2018 he was working as Assistant Professor in Bharti College of Engineering and Technology, Durg, Chhattisgarh, India. From June 2018 onward, he is pursuing Ph.D. from National Institute of Technology Silchar, Assam, India. His Research interests includes Power Quality, Grid Integration and application of renewable energy sources. He is a student member of IEEE.



Dr. Jyoti Prakash Mishra (M'09) received his Bachelor Degree (Electrical Engineering) from IE, India in 1996; M.Tech. Degree (Energy Management) from SEES, DAVV, Indore, India in Jan'1999 and Ph.D. Degree from IIT Roorkee, India in 2010. He worked as project-officer (Energy) for one year in NITCON, Chandigarh, India during 1999-2000; worked as Lecturer (Electrical) for more than two years in DIT, Dehradun, India during 2000-02 and then joined as a research scholar in IIT Roorkee. He joined NIT Silchar, Assam, India in Jan'2006 and is currently working as Associate Professor with the Department of Electrical Engineering. His research group bagged the IET Electric Power Applications Premium Award 2020 for one of their research paper. His Research interests includes Microgrid control (Grid Interactive/ Autonomous); Power electronics converters and its applications to Renewable Energy Systems; Power Quality, Power Management Issues and their Control, Multifunctional, Flexible Power Converters in distribution systems, Electric Vehicles applications. He is a member of IEEE & IE.



Dr. Amritesh Kumar (M'18) received the B.Tech. in Electrical and Electronics Engineering from SASTRA University, India in 2009, M.Tech. in Power Electronics and Drives from BITS PILANI, India in 2011 and Ph.D. from the Delhi Technological University (DTU), India in 2017. He is currently working as Assistant Professor with the Department of Electrical Engineering, NIT Silchar, India and heading the two SERB & one DST sponsored research projects. He possesses approximately nine years of teaching and research experience. He is recipient of premium research award (2019) at DTU and prestigious POSOCO award in 2018. His research interests include Analysis and design of power converters/inverters and its application in renewable energy sources, Multilevel systems, Power quality issues, AC–DC microgrids and Hybrid electric vehicles. He is a member of IEEE.

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:23:15 UTC from IEEE Xplore. Restrictions apply.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/352742100

Modulation of polymer-based immobilized enzymes for industrial scale applications

Chapter · January 2021

DOI: 10.1016/B978-0-12-819206-1.00003-X

CITATIONS
0

READS
94

2 authors:



Alka Dwevedi
Banaras Hindu University
43 PUBLICATIONS **523** CITATIONS
SEE PROFILE



Yogesh Sharma
Swami Shraddhanand College, University of Delhi
18 PUBLICATIONS **144** CITATIONS
SEE PROFILE

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

Preparation and Properties of Wax Emulsions, Wax Dispersion and Their Applications in Coatings View project

All content following this page was uploaded by Yogesh Sharma on 17 July 2021.

The user has requested enhancement of the downloaded file.

CHAPTER 3

Modulation of polymer-based immobilized enzymes for industrial scale applications

Alka Dwevedi[1] **and Yogesh K. Sharma**[2]
[1]Department of Biotechnology, Delhi Technological University, New Delhi, India
[2]Swami Shraddhanand College, University of Delhi, New Delhi, India

3.1 Introduction

Enzymes have become crucial part of various industries like development and production of nutritious food and beverage products, processing of fruit and vegetable, cheese, protein, grains, and fats and oils, as well as in several other industries like dairy, baking, brewing, and cereal extraction. Most importantly, enzyme-based processes result in higher yield, enhance the quality of final products, decrease costs, optimize resources, reduce wastage, and lower environmental pollution [1]. It has been seen since the last decade that enzymes have been sharing major part in global market with its increasing usage in various sectors. R&D centers of several companies have been now focusing on evaluation, development, upscaling (usually by expression and cloning of enzymes), validation of technologies, and innovative enzyme formulation for various commercial processes. Enzyme technology has been an integrated part of various interdisciplinary fields like biochemistry, organic chemistry, microbiology, and molecular biology. Enzyme technology has been known for several hundreds of years; however, 21st century has revolutionized it due to its increased applications in almost every sector [2]. It has been found that before 21st century, microbial enzymes (namely, bacterial, fungal, and yeast) have been most commonly used for various processes due to their easier purification. However, plant-based enzymes like amylase, protease, cellulase, and lipase have been now used most extensively due to the presence of evolved purification technologies [3]. Comparative analyses of enzyme sources have revealed that fungal enzymes have been most extensively used with respect to any other sources due to their easier growth and abundant enzyme production. Carbohydrases (cellulase, lactase, amylase, pectinase, and mannanases) have major contribution in industrial applications (namely, food and beverages, animal feed, textile, paper, biofuels, and pharmaceuticals) with major share in market growth [4] [Fig. 3.1 has summarized an impact of immobilized enzymes on global market (A), global market growth projections for technical enzymes by

Polymeric Supports for Enzyme Immobilization
DOI: https://doi.org/10.1016/B978-0-12-819206-1.00003-X
© 2021 Elsevier Inc.
All rights reserved.

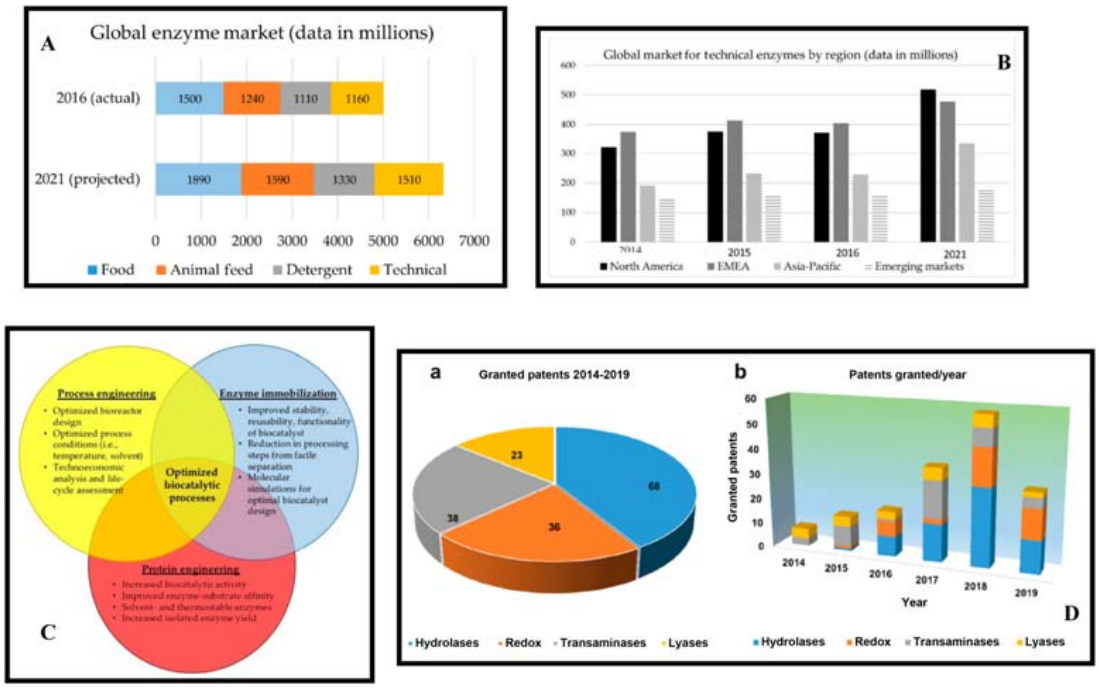


Figure 3.1 (A) Graphical representation of global enzyme market in 2016 (top) and projected global enzyme market in 2021 (bottom). (B) Global market growth projections for technical enzymes by geographical regions. (C) Schematic representation of a multidisciplinary approach aimed to define optimal biocatalytic processes. Implementation of biocatalysts in industrial technologies should have well established optimization protocols leading to increased enzyme operational stability at the interface with supports used for immobilization. (D) Patents granted in 2014–19 employing the most usual types of enzymes, according to several patent engine servers (e.g., Espacenet, SciFinder, Google Patents). *Adapted from (A) Dewan SS. Global markets for enzymes in industrial applications. Wellesley, MA: BCC Research; 2017; (B) and (C) Chapman J, Ismail AE, Dinu CZ. Industrial applications of enzymes: recent advances, techniques, and outlooks. Catalysts 2018;8:1–26; (D) de María PD, de Gonzalo G, Alcántara AR. Biocatalysis as useful tool in asymmetric synthesis: an assessment of recently granted patents (2014–2019). Catalysts 2019;9:1–42.*

geographical regions (B), optimization of industrial scale processes based on immobilized enzymes (C), and summary of granted patents (D)].

According to MarketsandMarkets, the enzymes have market of USD 10.6 billion (estimated for 2020), while it has been expected to reach USD 14.9 billion by 2026 according to Compound Annual Growth Rate (CAGR). Further, growing environmental concerns have propelled toward the production of clean energy that has resulted in contribution toward major share in the global market. It has been revealed that North America has major contribution followed by Asia Pacific, with key players across the world: BASF (Germany), DuPont (United States), Kerry Group (Ireland), Novozymes (Denmark), DSM (Netherlands), Chr. Hansen (Denmark), and Advanced Enzymes (India) (Fig. 3.2). The sectors like nutraceuticals and pharmaceuticals are majorly being developed in Asia Pacific



Figure 3.2 Industrial enzyme market in various countries across the world.

countries (like India and China) with priority focus on carbohydrases, while rest of the industrial applications have been attributed to North America and Europe. In the present decade, technological advancements have been made in bioethanol as well as in food and beverage industries. The applications of lipases (catalyzes hydrolysis of fats and lipids) have been growing with fast pace in the recent years due to their versatile applications in food industry (cheese and yogurt fermentation), baking, detergents, and production of biofuels [5]. Further, it has been widely used for biodiesel production that involves processing of glycerides and fatty acids via esterification and transesterification reactions in a cost-effective manner with minimal time and space consumption as well as easier modification and process optimization.

The industrial applications of enzymes have prerequisite of enzyme to be in its insoluble state as soluble state is fragile requiring extensive protocols for its stability during long storage as well as cost expensive due to its single use. The enzymes in its immobilized state can be used for several times with easier recovery from reaction mixture and maintained for large time periods. The enzyme immobilization is actually a platform that involves the combination of enzyme selectivity and kinetics with the physical and chemical properties of the immobilizing matrix in a specialized formulation to maximize both its stability and catalysis and make cost-effective industrial process [6]. It has been discussed in Chapter 1: Current and Future Trends on Polymer-Based Enzyme Immobilization, that the most commonly used immobilization methods are physical immobilization (adsorption or physical entrapment) and chemical immobilization (covalent binding and cross linking). It has been seen that sometimes whole cells containing enzymes are immobilized in various industrial processes to enhance the reusability cycles; however, it has often found to have lower catalytic efficiency with respect to isolated immobilized enzymes. Most importantly,

scaling up of immobilized process from lab to industries requires lesser optimization in a cost-effective manner than that based on soluble enzymes. Further, immobilized enzyme–based industrial processes have been enticing due to least time consumption and cost efficacious downstream processing. The industrial scale processes are based on bioreactors with several regulators for modulating physicochemical parameters to obtain product with highest yield [7]. Following parameters have been given utmost priority due to their direct correlation with product yield.

1. *Reaction temperature*: It has great impact on reaction kinetics. Generally, immobilized enzyme–based processes are operated at higher temperature ($\geq 10^{\circ}C$ with respect to soluble enzyme) due to their higher temperature stability.
2. *Amount of enzyme in immobilized state*: It is the major determinant for the overall cost of the production (generally 3%–10% w/v of immobilized enzyme/reaction total volume is being used in batch mode or fluidized mode run during industrial operations).
3. *Amount of substrate*: It becomes important when there is more than one substrate as they significantly affect the total product cost (usually substrates are taken in equimolar concentrations).
4. *Reusability cycle of immobilized enzyme*: It is extremely important to know exactly when to replace immobilized enzyme from the bioreactor as it affects the quality of product and efficacy of bioreactors (generally immobilized enzymes are replaced with in-between residual activity from 10% to 50% of the initial activity).

It is crucial to have thorough studies, especially with respect to technical details of the bioreactors. Bioreactors based on immobilized enzyme systems require rigorous studies on the shape and size of the immobilizing matrix. Here are given important

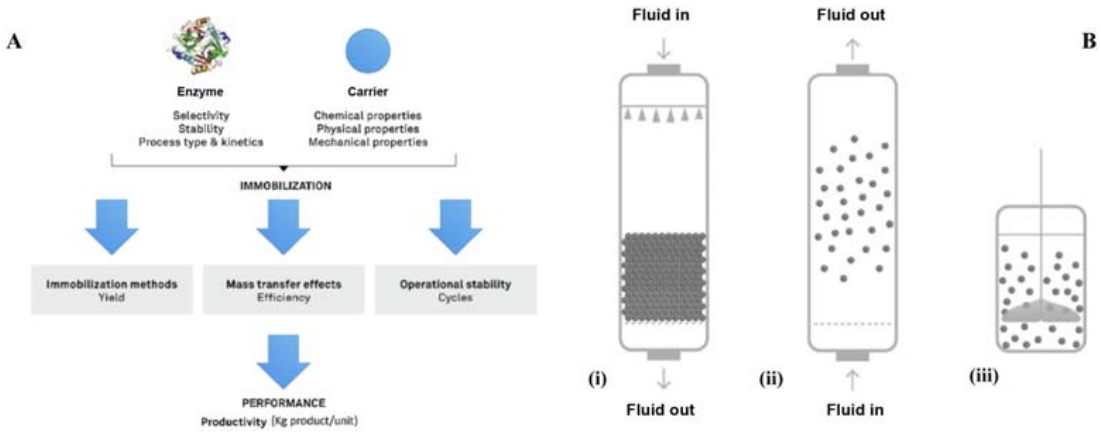


Figure 3.3 (A) Summary of important factors affecting enzyme immobilization. (B) Important reactors being used during industrial process based on immobilized enzyme (i) fixed bed reactor, (ii) fluidized bed reactor, and (iii) stirred reactor. *Adapted from Basso A, Serban S. Industrial applications of immobilized enzymes—a review. Mol Catal 2019;479:1–20.*

reactors being used during industrial process based on immobilized enzyme with their silent features [6] (Fig. 3.3).

Stirred batch reactors (commonly used in pharmaceutical industry):

1. Reactor volume is much smaller (i.e., few m^3) with respect to other known reactors. However, they have product yield in several hundred kilograms per batch.
2. Reactor contains stainless-steel filter at its bottom for easier liquid removal as well as for recycling immobilized enzyme. It contains a stirrer for controlled and gentle mixing of reaction mixture without affecting the mechanical stability of the immobilized enzyme. The size of the filter should be selected appropriately (generally lesser than the size of the immobilized enzyme) with high mechanical strength and allows easier cleaning.

Jacketed reactors (commonly used in food and chemical industries):

1. They have accurate control over temperature reaction parameter (especially important when immobilized enzyme has narrow optimum temperature).
2. They have smaller internal volume of $\geq 10\ m^3$ with productivity of over ton.
3. Generally operate continuously with continuous monitoring of inline controls.

Fluidized bed reactor:

1. A best option when there are substrates and products with high viscosity.
2. Both immobilized enzymes and substrates are in suspension with appropriate flow rate of substrate.
3. All the physicochemical parameters are continuously monitored.

3.2 Industrial applications of enzymes

The enzyme catalysis has been successfully implemented in almost all industries, with wider usage in pharmaceuticals, food and beverage, detergent, biofuel, etc. having good annual turn overs. Chemical-based industries have been a major contributor toward nations' economy; however, they have major setbacks on environment. There are continuous efforts being made to replace most of the chemical industries with enzyme-based industries. Table 3.1 has provided a summary of industries based on enzyme catalysis (immobilized states with polymeric matrices). It has been anticipated that global enzyme market would show an increase by 30% in 2021 as compared to 2016, according to research by BCC (Business Communication Company) on various industries like textile, paper, leather, and biodiesel. Further, enzyme immobilization would provide an ideal platform for widening enzyme-based industries as it helps in scaling up the process by several folds with least consumption of man power in a cost-effective manner. Following section has provided an overview on enzyme-based industries.

Table 3.1 Industrial applications of enzyme catalysis.

Sector	Enzymes	Applications
Pharmaceuticals	Nitrile hydratase, transaminase, monoamine oxidase, lipase, penicillin acylase	Synthesis of intermediates for the production of active pharmaceuticals ingredients
Food processing	Trypsin, amylase, glucose isomerase, papain, pectinase	Conversion of starch to glucose, production of prebiotics, debittering of fruit juice
Detergent	Protease, lipase, amylase, cellulase	Stain removal, removal of fats and oils, color retention
Biofuels	Lipase, cellulase, xylanase	Production of fatty acid methyl esters, decomposition of lignocellulotic material for bioethanol production
Paper and pulp	Lipase, cellulase, xylanase	Removal of lignin for improved bleaching, improvement in fiber properties

Source: Adapted from Chapman J, Ismail AE, Dinu CZ. Industrial applications of enzymes: recent advances, techniques, and outlooks. Catalysts 2018;8:1–26.

3.2.1 Pharmaceutical industry

Enzyme catalysis has been an utmost priority by pharmaceutical industries due to its high regio-, chemo-, and stereoselectivity (often lacking in chemical-based catalysis) in the generated product at industrial scale [8] [Table 3.2 has provided large-scale synthesis of important API (active pharmaceutical ingredients) based on immobilized enzymes] (Basso and Serban, 2019). It is extremely important to have high degree of product specificity during pharmaceutical processes to improve process economics and avoid any health side effects (due to any small amount unwanted product). Enzyme-based pharmaceutical processes do not require extreme physicochemical conditions like harsh chemicals and high temperature and produce higher yield of specific enantiopure product, thus making the process safe for human applications (Fig. 3.4 has shown summary of patented and commercialized pharmaceuticals based on enzyme catalysis). Following section has provided successful example of enzyme-based pharmaceutical processes at commercial scale:

1. *Sitagliptin*: It is used for the treatment of type II diabetes, being marketed by the company Merck. It is a dipeptidyl peptidase-4 inhibitor which helps in the prevention of diabetes-induced tight junction disassembly and tightens blood–retinal barrier [9]. The chemical synthesis of sitagliptin has been not very successful as it requires very high pressure for the hydrogenation of enamine in the presence of rhodium. Further, it leads to the production of nonspecific enantiomers (sometimes goes to 97%). There is much more tedious downstream processing that requires removal of trace amount of

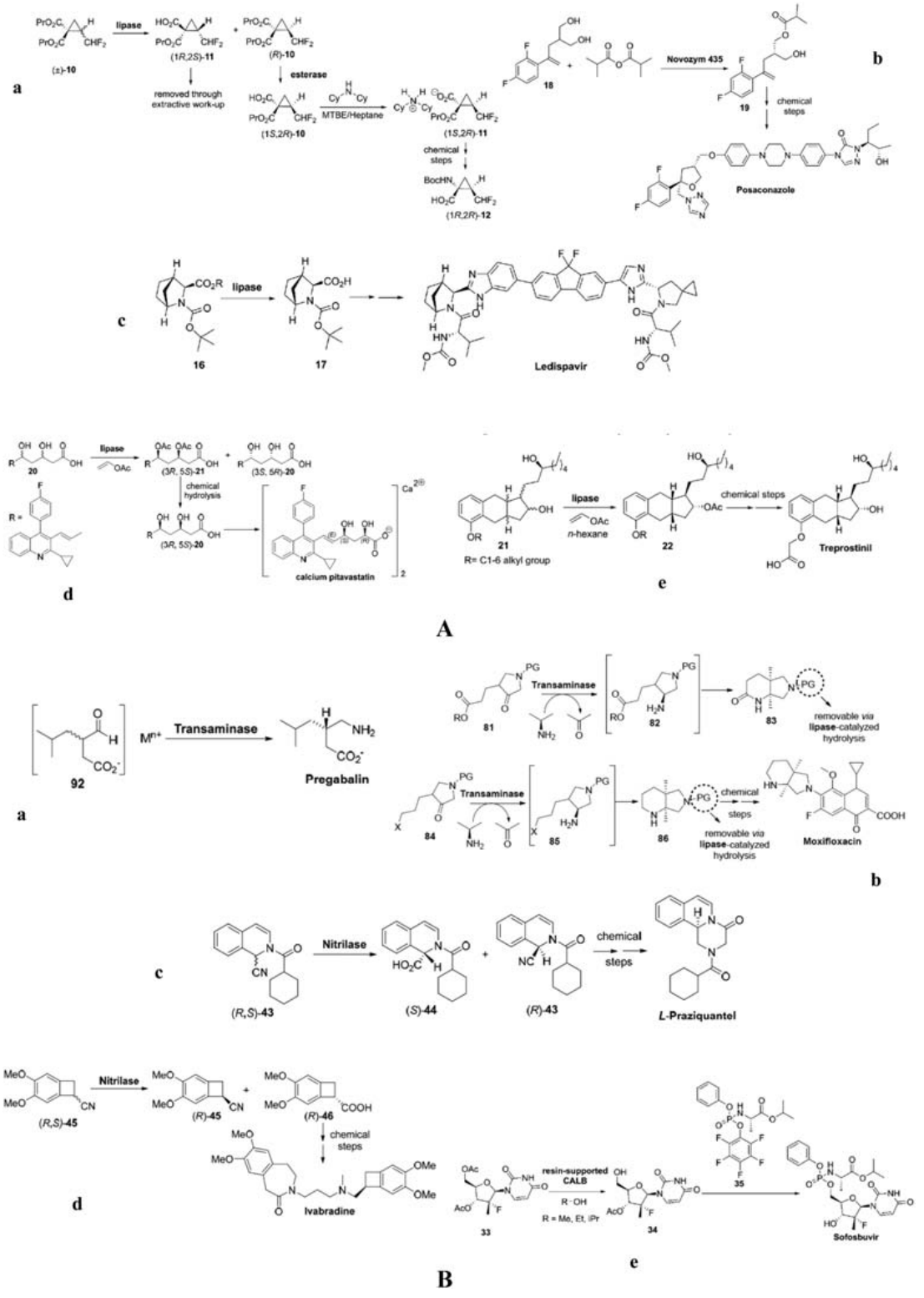
Table 3.2 Industrial scale synthesis of commercially important active pharmaceutical ingredients using immobilized enzymes (hydrolases and transferases).

Active pharmaceutical ingredients	Target applications	Immobilized enzyme	Type of immobilization	Reaction type
Odanacatib	Osteoporosis treatment	Lipase B from *Candida antarctica*	Adsorption on octadecyl polymethacrylate resin	Esterification
Sofosbuvir	Treatment of hepatitis C	Lipase B from *C. antarctica*	Adsorption on methacrylate/ divinylbenzene copolymer	Hydrolysis
6-APA	Beta-lactams	Penicillin amidase (hyd)	Covalent on epoxy or amino methacrylate polymer	Hydrolysis
Amoxicillin/ ampicillin	Beta-lactams	Penicillin amidase (syn)	Covalent on epoxy or amino methacrylate polymer	Amidation
Sitagliptin	Antidiabetic	Transaminase	Adsorption	Transamination

6-APA, 6-Aminopenicillanic acid.
Source: Adapted from Basso A, Serban S. Industrial applications of immobilized enzymes—a review. Mol Catal 2019;479:1–20.

rhodium and isolation of required enantiopure product to be fitted for human usage. Therefore enzyme-based process has been opted (by Codexis and Merck) requiring R-selective transaminase (R-ATA) from *Arthrobacter* sp. with the production yield of 200 g L^{-1} by converting prositagliptin ketone to sitagliptin (in the presence of dimethylsulfoxide). The purity of obtained sitagliptin is more than 99.0% without requirement of any heavy metal or extreme physicochemical conditions [10].

2. *Boceprevir*: It is an NS3 protease inhibitor used for the treatment of chronic hepatitis C. It is being marketed by Codexis and Merck, based on monoamine oxidase (MAO)-catalyzed process leading to the production of enantiopure bicyclic proline intermediate, which is a precursor of boceprevir. The protein engineering has been very useful in producing recombinant MAO with excellent catalytic activity, thermostability, and good solubility in the reaction mixture making the process compatible for industrial setup at large scale. Conventional chemical synthesis of bicyclic proline (precursor of boceprevir) requires metal-based oxidant and reductant. It is a labor-intensive process requiring almost eight steps without assuring product purity with the generation of tremendous amount of


lipase
esterase
removed through extractive work-up
MTBE/Heptane
chemical steps
Novozym 435
Posaconazole
Ledispavir
R= C1-6 alkyl group
n-hexane
Treprostinil
calcium pitavastatin
Transaminase
Pregabalin
removable via lipase-catalyzed hydrolysis
Moxifloxacin
Nitrilase
L-Praziquantel
Ivabradine
resin-supported CALB
R = Me, Et, iPr
Sofosbuvir
A
B

(*Continued*)

nonecofriendly waste. The enzyme-catalyzed process requires a lesser amount of raw material and water ranging between 40.2% and 67.2%, respectively, with respect to chemical-based synthesis with the generation of very less of amount of waste (nontoxic). Further, it is economically more efficient requiring least labor and time consumption [11].

3. *(S,S)-Reboxetine succinate*: It is a noradrenergic antidepressant used for the treatment of fibromyalgia. It is being synthesized and marketed by Pfizer at commercial scale based on lipase (from *Candida antarctica*) catalysis. In either case, enzyme- or chemical-based synthesis of (*S*,*S*)-reboxetine succinate requires an acetylation of diol intermediate. Enzyme-based synthesis has regioselectivity of 98% with a yield of $\geq$99%. However, classical chemical acetylation of diol intermediate suffers from poor enantioselectivity and generation of lots of unwanted byproducts. Further, enzyme-based synthesis is applied at lower cost of production with least time consumption [12].
4. *Pregabalin*: The industrial scale production of pregabalin (neuroactive drug) is used in pain killing, antianxiety, and as an anticonvulsant. Thus it is useful for the treatment of epilepsy, anxiety, and social phobia. The enzyme lipolase (commercially available lipase) is used in its synthesis as well as separation of S-enantiomer from R-enantiomer intermediate. It requires divalents like Ca^{2+} and Zn^{2+}, which has excellent selectivity and cost-effectiveness, and generates less waste (just 20% of product). The obtained product has a purity of 99.5% with yield of 40%–45% higher than classical chemical synthesis route [13].

3.2.2 Fuel-based industries

The broad categories of biofuels being synthesized by enzymes are hydrogen, ethanol, and biodiesel derived from biomass, animal fats, waste oils, etc. They have been good potential to substitute fossil fuels to meet global energy demand with excellent efficiency and cost efficacy. Further, these fuels are environmentally friendly due to lower

◀ **Figure 3.4** Summary of patented and commercialized pharmaceuticals based on immobilized enzymes. (A) (a) Chemoenzymatic synthesis of an enantiopure precursor of a viral protease inhibitor; (b) chemoenzymatic preparation of an intermediate in the synthesis of posaconazole; (c) enzymatic preparation of a ledipasvir intermediate; (d) preparation of pitavastatin through an enzymatic transesterification; and (e) preparation of treprostinil through an enzymatic acylation. (B) (a) Transaminase catalyzed step for the synthesis of pregabalin, patented by Pfizer; (b) use of transaminases to afford 3-aminopyrrolidines, as precursors of pharmaceutical compounds, for example, moxifloxacin; (c) use of nitrilases in the synthesis of L-praziquantel; (d) synthesis of precursors of the drug ivabradine using nitrilases; and (e) chemoenzymatic synthesis of sofosbuvir. *Adapted from de María PD, de Gonzalo G, Alcántara AR. Biocatalysis as useful tool in asymmetric synthesis: an assessment of recently granted patents (2014–2019). Catalysts 2019;9:1–42.*

emission of toxic gases like carbon monoxide, carbon dioxide, nitrogen oxides, sulfur oxides, and particulate matter. The reports by various agencies have revealed that the global production of biofuels has increased in the last decade by over 20% and the trend is still going on, which indicates that coming years would come up with several automobiles running on biofuels. It has provided platform to revolutionize agriculture and provide several folds higher turnover as well as direct linkage to industries [7]. The following section has provided an overview on successful available biofuels derived from enzyme catalysis.

1. *Biodiesel*: It is based on esterification of fatty acids (derived from waste products like vegetable oil and animal fat) in the presence of alkali. It is also termed FAMEs (fatty acids methyl esters) produced from fatty acids in the presence of appropriate catalysts. The conventional chemical process is based on conversion of fatty acids derived from plants or animals into FAMEs by using sodium methoxide followed by product purification contaminated with high alkali salt, often a tedious and costly process. The enzyme-based process applies lipase for esterification with excellent efficiency under mild reaction conditions leading to high-purity product yields with minimal requirement of product purification making the complete process cost-effective and least laborious. Lipases are the excellent biocatalysts due to their remarkable biochemical and physiological properties and have been most desirable by industries due to its implications in various processes like alcoholysis, acidolysis, aminolysis, and hydrolysis reactions [7]. The enzyme is available from broad variety of sources like:

1. *Plants*: papaya latex, oat seed lipase, and castor seed
2. *Animals*: pig, human pancreas
3. *Bacteria*: *Chromobacterium viscosum*, *Photobacterium lipolyticum*, *Pseudomonas cepacia*, *Pseudomonas fluorescens*, *Streptomyces* sp., and *Thermomyces lanuginose*
4. *Filamentous fungi*: *Aspergillus niger*, *Mucor miehei*, and *Rhizopus oryzae*
5. *Yeast*: *C. antarctica* and *Candida rugosa*

The production of biodiesel by using lipase (first reported by Mittelbach) follows ping-pong bi−bi mechanism during fatty acid trans-esterification involving hydrolysis of ester bond followed by esterification with the second substrate [14]. It has been reported that most of the processes for biodiesel production are based on lipase from yeast due to its higher product yield, least growth requirement, shortest generation time, excellent adaptability toward extreme environmental conditions, and simplicity in genetic manipulation. There are three broad categories of lipases based on their specificity: 1,3-specific lipases, fatty acid−specific lipases, and nonspecific lipases. Most of the industrial processes have reported that 1,3-specific lipases are best for the production of biodiesel due to its catalytic efficiency and excellent specificity. Among companies, Hainabaichuan Co. Ltd. (Guangzhou, China) has reported on production with an annual capacity of 20,000 t in 2006 and 40,000 t in 2008 based on

lipase-catalyzed process using waste palm oil and waste edible oil, while Lvming Environmental Technology Co. Ltd. (Shanghai, China) has reported an annual capacity of 10,000 t in 2007 using immobilized lipase, and Piedmont Biofuels has reported on a production of biodiesel with an annual capacity of 15,000 t in 2012 based on continuous enzymatic trans-esterification of free fatty acids by immobilized lipase from *C. antarctica* lipase. Immobilization of lipase has been useful for lowering cost of biodiesel production by several hundred folds; however, it often leads to lower catalytic efficiency of lipase.

2. *Bioethanol*: It is chemically ethyl alcohol, now being used as fuel (namely, as an additive for gasoline in motor fuel). It is derived from various agricultural feedstocks like corn, sugarcane, hemp, potato, and cassava. However, it leads to negative impact on food prices by increasing their cost as well as requirement of large amount of cultivable land for growing crops. Therefore an alternative substitute has been looked to be used as source for bioethanol production. Cellulosic, lignocellulosic (contain lignin, cellulose, and hemicellulose) biomasses and starch grains are found in abundance in the cell wall of plants. Further, municipal solid waste can be used in addition to previous sources for the production of ethanol in substantial amount. There are two ways of producing ethanol from these sources: chemical and enzymatic. Chemical-based method requires harsh chemicals, high temperature ($\sim 100^{\circ}C$) and extremely low vapor pressure for sample treatment, and processing to release glucose followed by its subsequent conversion into bioethanol via whole-cell fermentation. Therefore implies high energy inputs and the generation of large amount of toxic byproducts. In the case of enzyme-based bioethanol generation, it requires enzymes like cellulase, invertase, β-glucosidase (BGL), and xylanase for biomass pretreatment and hydrolyzation into fermentable sugar followed by ethanol production [7]. It is least time consumptive (due to excellent substrate specificity), cost-effective, requires mild reaction conditions, and generates a minimal amount of waste (nontoxic). It has been reported that worldwide production of bioethanol has increased by five times during last decades; the United States and Brazil are known to be the top most producers. It has been recommended by many of the countries to use bioethanol as fuel in its blended state (usually E10, 10% ethanol) in most of vehicles manufactured after 2000. It has been reported that immobilized states of enzymes (cellulase, invertase, BGL, and xylanase) are most preferred for industrial scale synthesis of bioethanol. The following section has provided an overview of various matrices (leading to improved physicochemical properties of enzymes) being used for their immobilization.

1. Cellulases (mostly used from *Trichoderma reesei*)
   a. *Chitosan-coated magnetic nanoparticles* [15]
      i. Covalent binding by using glutaraldehyde.

ii. Average diameter of beads: 8–10 nm.
iii. K_m value has been increased by eight times with respect to soluble enzyme.
iv. Reusability cycles: 15.
v. Better thermostability with respect to soluble enzyme.
vi. Longer storage stability.

b. *Supermagnetic $Fe_3O_4 \cdot SiO_2$ nanoparticles* [16]
i. They have core–shell structure with diameter of 10 nm.
ii. Enzyme attachment via adsorption.
iii. Immobilization efficiency of 95%.
iv. Shelf half-life has been increased by 3.3 folds with respect to soluble enzyme at 70°C.
v. No change in optimal pH.
vi. Better thermostability.
vii. Higher catalytic efficiency.
viii. Besides being used for bioethanol production, it is also used in paper, pulp, and pharmaceutical industries.

c. *PEGylated graphene oxide (GO) nanosheets* [17]
i. Stability of immobilized enzyme has increased by 30 folds.
ii. K_m of immobilized enzyme is 3.2 mg mL^{-1}, while that of soluble enzyme is 2.7 mg mL^{-1}.

d. *Magnetic cross-linked enzyme aggregates (CLEAs)* [18]
i. Immobilized enzyme has wider optimum pH and temperature.
ii. Reusability cycles: 6, with retention of 74% initial activity.
iii. Bioethanol yield of 21% has been obtained using bamboo biomass.

e. *Superparamagnetic FeO nanoparticles* [19]
i. Binding capacity of matrix: 0.43 g of enzyme per gram of matrix.
ii. Reusability cycles: 10, with retention of 75% of initial activity.
iii. Better storage stability with respect to soluble enzyme.
iv. Susceptibility toward aggregation, especially in air.

f. *MgO nanoparticles with surface layer of GO* [20]
i. Cellulase from *Bacillus subtilis* (psychrophilic strain).
ii. Enzyme immobilization by using glutaraldehyde.
iii. Immobilized enzyme has higher enzymatic activity (~2.98 folds) with respect to soluble enzyme at 8°C.
iv. K_m of immobilized enzyme has been reduced by 6.7 folds, while V_{max} has been increased by 5 folds with respect to soluble enzyme at 8°C.
v. Reusability cycles: 12.
vi. Storage stability of 120 days at 4°C.
vii. Higher thermostability can work even at 90°C without much loss in enzyme activity.

g. *TiO_2 nanoparticles* [21]
- **i.** Cellulase from *Aspergillus niger*.
- **ii.** Enzyme immobilization by physical adsorption and covalent coupling (using aminopropyltriethoxysilane) with immobilization efficiency of 76% and 93%, respectively.
- **iii.** Catalytic efficiency (V_{max}/K_m) has been increased by 10 times during covalent binding of enzyme, while it has been increased only by 3 times during enzyme binding by adsorption.
- **iv.** Covalently attached enzyme has better thermal and operational stability, while enzyme attached by adsorption has better reusability with respect to soluble enzyme.

2. Invertase

a. *Magnetic diatomaceous earth nanoparticles* [22]
- **i.** Immobilization efficiency of 92.5%.
- **ii.** Specific activity of immobilized enzyme has been increased by 2.42 folds with respect to soluble enzyme.
- **iii.** Better thermostability with retention of 85% of enzyme activity after 60 min at 35°C.
- **iv.** Better storage stability with respect to soluble enzyme with retention of 80% enzyme activity after 120 days at 4°C.

b. *Magnetic nanoparticles (Fe_3O_4), modified with silica ($Fe_3O_4 \cdot SiO_2$)* [23]
- **i.** Enzyme has been attached to it after addition of GMA (glycidyl methacrylate) followed by the addition of spacer arm hexamethylene diamine.
- **ii.** Amount of enzyme immobilized: 33.4 ± 1.3 mg g^{-1}.
- **iii.** K_m and V_{max} of immobilized enzyme: 39.4 mmol L^{-1} and 349.5 mmol L^{-1} min^{-1} (K_m and V_{max} of soluble enzyme are 34.3 mmol L^{-1} and 387.2 mmol L^{-1} min^{-1}, respectively).

c. *Chitosan-coated sol–gel-derived γ-Fe_2O_3 magnetic nanoparticles* [24]
- **i.** Reusability cycles: 20, almost no loss in enzyme activity.
- **ii.** Immobilized enzyme is stable over wide pH and temperature.
- **iii.** Metal divalents like Cu^{2+}, Zn^{2+}, Cr^{2+}, and Ni^{2+} are used for enzyme adsorption; however, maximum immobilization efficiency has been attached in the case of Cu^{2+} at pH 5.0.

d. *Chitosan nanoparticles* [25]
- **i.** Invertase taken from *Saccharomyces cerevisiae*.
- **ii.** Immobilizing matrix has been activated with glutaraldehyde.
- **iii.** Enzyme immobilization by covalent linkage.
- **iv.** Physicochemical parameters like thermal and storage stabilities, optimal pH, and temperature of immobilized enzyme are not changed with respect to soluble enzyme.

v. Reusability cycles: 59, almost no loss in enzyme activity.
vi. Excellent operational stability with respect to various published reports.

3. BGL [26]
   e. *SiO_2 nanoparticles*
      i. Immobilization efficiency of 52% has been achieved with a yield of 14.1%.
      ii. Optimal temperature and pH of immobilized enzyme are 60°C and 5.0.
      iii. Thermostability of immobilized enzyme: 60°C–70°C.
      iv. K_m and V_{max} of immobilized enzyme: 1.074 mM and 1.513 U mg^{-1} (soluble enzyme: 0.9 mM and 3.5 U mg^{-1}).
      v. Improved thermo and storage stabilities.
      vi. Reusability cycles: 10, with retention of 70% of initial enzyme activity.
      vii. Alternative source of BGL are sugarcane juice, caffeic acid, etc.
   f. *Magnetic Fe_3O_4 nanoparticles (MNPs)* [27]
      i. Recombinant BGL has been used by expressing its gene taken from *Coptotermes formosanus* in *Escherichia coli* BL21.
      ii. Requires metal divalents like Cu^{2+}, Zn^{2+}, Cr^{2+}, Ni^{2+}, and Co^{2+} for BGL immobilization, with maximum immobilization efficiency that has been attached in the case of Co^{2+}.
      iii. K_m and V_{max} of the immobilized enzyme: 0.904 mM and 0.057 μmol min^{-1}.
      iv. Immobilized enzyme has higher thermostability and operational stability.
      v. Reusability cycles: 15, with retention of >90% of its initial enzyme activity.
   g. *Fe_3O_4/PMG/IDA-Ni^{2+} nanoparticles* [28]
      i. They are chemically Fe_3O_4/PMG [poly(*N*,*N*′-methylenebisacrylamide-*co*-GMA)] core/shell microspheres with iminodiacetic acid to open epoxy rings present on matrix shell to be used for covalent enzyme attachment.
      ii. Immobilized enzyme has excellent catalytic activity and operational stabilities.
      iii. Reusability cycles: 30, with retention of 65% of initial enzyme activity.
   h. *Magnetic nanoparticles* [29]
      i. Thermostable BGL isolated from *A. niger*.
      ii. Enzyme immobilization via covalent binding.
      iii. Immobilization efficiency: 93%.
      iv. Temperature and pH optima of immobilized enzyme: 60°C (similar to soluble enzyme), 6.0 (soluble enzyme has 4.0).
      v. K_m of immobilized and soluble enzymes: 3.5 and 4.3 mM, respectively.
      vi. Reusability cycles: 16, with retention of 50% of initial enzyme activity.
      vii. Immobilized enzyme has improved thermostability at 70°C.
   i. *Silicon oxide nanoparticles* [30]

i. Crude enzyme preparation of BGL has been used for immobilization, isolated from *Agaricus arvensis*.
ii. Enzyme attachment via covalent binding.
iii. Immobilization efficiency: 158%.
iv. Immobilized enzyme has improved thermostability with half-life increased by 288 folds at 65°C.
v. Catalytic properties of immobilized enzyme: V_{max} (3028 U mg protein^{-1}) and k_{cat} (4945 s^{-1}), while that of soluble enzyme with V_{max} (3347 U mg protein^{-1}) and k_{cat} (5466 s^{-1}).
vi. Reusability cycles: 25, with retention of 95% of initial enzyme activity.

j. *Single-walled carbon nanotubes incorporated with Fe_2O_3* [31]
i. Enzyme immobilization by both physical adsorption and covalent binding.
ii. Immobilization efficiency: 40%.
iii. Reusability cycles: 10, with almost no loss in enzyme activity.
iv. Storage stability: 1 month at 4°C.
v. Best suited for industrial scale processes (namely, large-scale bioreactors).

4. Xylanase from *Armillaria gemina* [32]

k. *Functionalized silicon oxide nanoparticles*
i. Catalytic efficiency ($k_{cat}/K_m = 1440$ mg mL^{-1} s^{-1}).
ii. Immobilization efficiency of 117%.
iii. There is slight shift in optimal pH and temperature with respect to soluble enzyme.
iv. Shelf half-life has been increased by four folds.
v. Reusability cycles: 17, with retention of 92% of initial enzyme activity.

l. *Magnetic nanoparticles modified with hyperbranched polyglycerol* [33]
i. Xylanase from *Pholiota adipose*.
ii. Immobilized xylanase has shown substantial increased in reactivity, reusability, and stability.
iii. Best suited for large-scale industrial processes.
iv. k_{cat} of 4261 s^{-1}, much higher than soluble enzyme.

m. *Silicon oxide nanoparticles* [34]
i. Immobilization efficiency of 66%.
ii. Reusability cycles: 17, with retention of 97% of initial enzyme activity.
iii. Well suited for large-scale industrial processes.

3. *Extraction and recovery of natural gas*: There are vast reserves of natural gas across the world, with the United States having 8.7 trillion m^3 (20% of global natural gas in 2016). It has been estimated that there is 80%–95% of methane with varying amount of hydrocarbons in natural gas, creating much hurdles during storage and transportation of natural gas that limits its usage as fuel. Further, methane has been very useful for various industrial processes in addition to its application as fuel.

Technological developments have motivated to convert methane into methanol that can be readily converted into olefins and other valuable hydrocarbons. Both chemical and enzymatic methods are being used; the former is based on Fischer–Tropsch process with requisite of high temperature, high pressure in addition to noble metal catalysts. However, chemical method has lower yield with poor methanol recovery, cost expensive, and much laborious. In the case of enzymatic method, methane monooxygenase is used to convert methane into methanol under mild conditions with yield $\geq 99.0\%$ in short-time duration in a cost-effective manner with least labor [35]. It has been used for large-scale production of methanol from methane by using three-dimensional (3D)-printed microbioreactor. Further, immobilized enzyme is more useful particularly in lowering total cost of methanol production by several folds with easier recovery of immobilized enzyme after every reaction cycle. However, there is lowering of catalytic efficiency of enzyme due to immobilization particularly due to changed enzyme conformation and due to mass transfer limitations (due to lower solubility of oxygen and methanol in aqueous media). Later can be rectified by much extent through working on process configurations. Therefore collectively enzyme-based process in the production of methanol has been industrially much preferred and easily adopted even at site of natural gas extraction.

3.2.3 Food and beverage industries

This industry completely relies on enzyme-based products as traditional chemical synthesis may lead to various complex reactions, reagent toxicity and is not economically viable [5] (Table 3.3 has provided commercially available food ingredients synthesized using immobilized enzymes). Following bullet points have provided enzyme-based applications in food and beverage industries:

1. Wine and cheese making (enzyme-based process known for back thousand years).
2. Food preservation and modification (namely, beer fermentation, juice debittering, and bread baking).
3. Conversion of starch into glucose and fructose by using enzymes like α-amylase at 100°C for sequential liquefaction and saccharification, known several decades ago (alternative chemical methods require 175°C and provide no surety over food toxicity).
4. Production of fructose and maltose syrup.
5. Commercial scale production of probiotics (like oligosaccharides, lactulose, lactitol hydrolysates, and inulin as nondigestible food additive which help to stimulate growth of gut bacteria and improves human health), artificial sweeteners, and rare sugars. Companies like Yakult Honsha Co. Ltd. Of Japan and Friesland Food Domo of The Netherlands are producing GOS (galacto-oligosaccharides) at

Table 3.3 Implementation of immobilized enzymes for large-scale synthesis of food ingredients.

Product	Application	Immobilized enzyme	Reaction type	Type of immobilization
HFCS	Sweetener	D-Glucose/xylose isomerase	Isomerization	Adsorption on silica
Allulose	Sweetener	3-Epimerase	Epimerization	Adsorption/ionic
Tagatose	Sweetener	Beta-galactosidase, epimerization by baker yeast	Hydrolysis	Adsorption/ionic ion exchange styrene/Divinyl Benzene (DVB) polymer
GOS	Synthesis of oligosaccharides	Beta-galactosidase	Synthesis	Covalent immobilization on synthetic polymers
Vitamin C ester	Antioxidant	Lipase B from *Candida antarctica*	Transesterification	Adsorption on methacrylate/ divinylbenzene copolymer
Triglycerides	Cocoa butter equivalent	Lipase from *Thermomyces lanuginosus*	Transesterification	Adsorption on silica
Triglycerides	Cocoa butter equivalent	Lipase from *Rhizomucor miehei*	Transesterification	Adsorption/ionic on ion exchange styrene/DVB polymer
Omega 3 ethyl esters	Omega 3 from fish oil	Lipase B from *C. antarctica*	Esterification	Adsorption on methacrylate divinylbenzene copolymer
Hydrolyzed lactose	Lactose free dairy products	Beta-galactosidase	Hydrolysis	Adsorption/ionic on ion exchange—modified cellulose fiber

GOS, Galacto-oligosaccharides; *HFCS*, high fructose corn syrup.
Source: Adapted from Basso A, Serban S. Industrial applications of immobilized enzymes—a review. Mol Catal 2019;479:1–20.

commercial scales through transgalactosylation of lactose by using β-galactosidase (either soluble or immobilized based on optimized process).

6. Proteolytic enzymes (namely, papain) are used for producing protein hydrosylates for applications in nutritional supplements, flavor enhancers, and therapeutics (antioxidant, antimicrobial, and antihypertensive). Further, the production of fish protein

hydrolysates by processing of seafood helps in minimizing pollution from fishing industries; protein hydrolysates from Chinese walnuts have excellent antioxidant properties, etc.

3.2.4 Flavors and aromas-based industries

These industries have been well dominated by chemical-based process due to its versatility; however, there are reports mentioning implementation of enzyme-based catalysis. The supporting points include: it is much more economically beneficial, ecofriendly, and allows easier scaling up due its regio-, chemo- and stereoselectivity (helps in accurate positioning of functional groups like alcohols, aldehydes, ketones, and esters in the aromatic compounds). Enzyme-based processes have anticipated to alleviate annual turnover of global market based on flavors and fragrances. The enzyme lipase has been thoroughly explored in flavors and aromas-based industries due to its capability to transfer acyl groups from esters to other nucleophiles [5]. Further, protein engineering, process engineering, and enzyme immobilization have been very useful in facilitating profitable industrial processes. Few of the examples of industrial processes have been given here.

1. The implementation of enzyme lipase to convert oleic acid to benzyl oleate with benzyl alcohol as an acyl acceptor (achieved efficiency of 94%) [36].
2. The implementation of immobilized lipase for converting vinyl propionate to *p*-cresyl propionate in the presence of heptane (achieved efficiency of 99%) [37].

3.2.5 Detergent industry

It is well known that discharged, unused chemical detergents in water bodies have been major nuisance to environment due to resistance toward degradation, thus leading to enhanced Biochemical Oxygen Demand (BOD) and Chemical Oxygen Demand (COD) of water bodies. Further, they poses higher health risks as their complete removal by various water purifiers is still not 100% guaranteed and, thus gets accumulated inside living system for very long duration due to their resistance toward degradation by available digestive enzymes. It has been essentially crucial to replace chemical-based detergents with enzymes as the latter is ecofriendly, mitigates least damage to fabrics and surfaces, and much more cost-effective. It has been researched by various scientific communities across the world to look for various enzymes that can replace chemical detergents [1]. Till date, the amount of chemical detergents can be lowered by addition of various enzymes like α-amylase, lipase, and alkaline protease. However, the complete replacement of chemical detergents is still a challenge.

1. The enzyme alkaline protease is effective in the removal of protein stains as well as in cleaving damaged cotton fibers. It is usually isolated from microbial sources for commercial process. The enzyme has been used on commercial scale by various renowned companies like Novozymes SA, Kao Corporation, Genencor International.

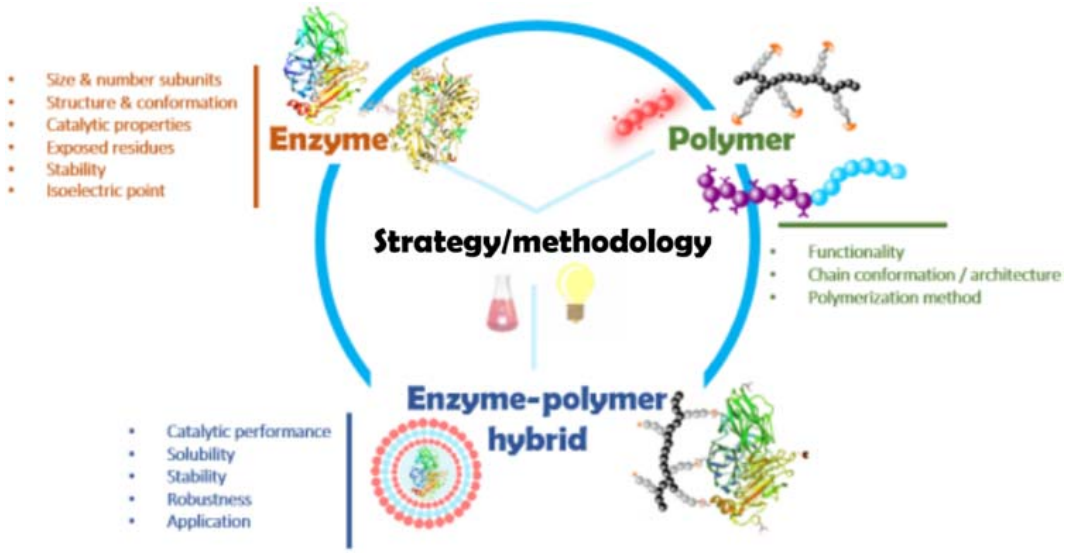


Figure 3.5 Optimization of polymer-based enzyme immobilization to be used for industrial scale applications. *Adapted from Rodriguez-Abetxuko A, Sánchez-deAlcazar D, Muñumer P, Beloqui A. Tunable polymeric scaffolds for enzyme immobilization. Front Bioeng Biotechnol 2020;8:1–27.*

2. The enzymes like amylases and lipases in varied ratios based on their process optimization are used for various applications like in dishwashing detergents for the removal of starch food deposits, fats, and oils.

3.3 Modulating polymers for industrial applications based on immobilized enzymes

Enzyme immobilization has been known for more than 100 years. It is the simplest technique known for modulating enzymes according to the wide range of industrial applications (Fig. 3.5). Immobilized enzymes are best suited for all types of enzyme-based catalysis involved in the synthesis of various fine chemicals, intermediates, pharmaceuticals, etc. under mild conditions. Genetic engineering has been opted for modifying enzymes to overcome various limitations being generally confronted during large-scale industrial processes like operational and storage stabilities, resistance toward extreme physicochemical conditions, catalytic efficiency, etc. However, it has not been as successful as predicted due to various challenges like requirement of lots of expertise, time consumptive, cost expensive, posttranslational modifications are not done (helpful for effective enzyme functioning in various cases), etc. Therefore enzyme immobilization is the only best option for large-scale industrial processes based on enzyme catalysis with an additional advantage of enzyme reusability and easier separation from reaction mixture (minimizes downstream processing that makes the process cost-effective, reliable, and efficient). The source of enzyme has been significant as it decides the amount of enzyme yield, natural enzyme stability, and

catalytic efficiency [38]. Most commonly, microbial source of enzyme is used as it has the highest yield with respect to animal or plant source. However, catalytic efficiency of enzymes is usually higher in the case of plant or animal source due to their higher order of complexity. Further eukaryotic enzymes are highly regio- and stereospecific with ability to produce chirally pure products (not easily possible by chemical synthesis or microbial enzymes). Thus they are very useful especially in pharmaceuticals and biofuels by carrying out effective reactions like reductions, oxidations, hydroxylations, acetylations, epoxidations, glycosylations, (de)esterifications, (de)methylations, and isomerizations.

Enzyme immobilization has been useful in promoting physicochemical properties (high activity, selectivity, operational stability, productivity, etc.); however, it may also lead to severe loss of catalytic efficiency due to various reasons. However, there is no generic method that can provide accurate guidelines for enzyme immobilization. It has been still not established that type of immobilization method to be adopted for

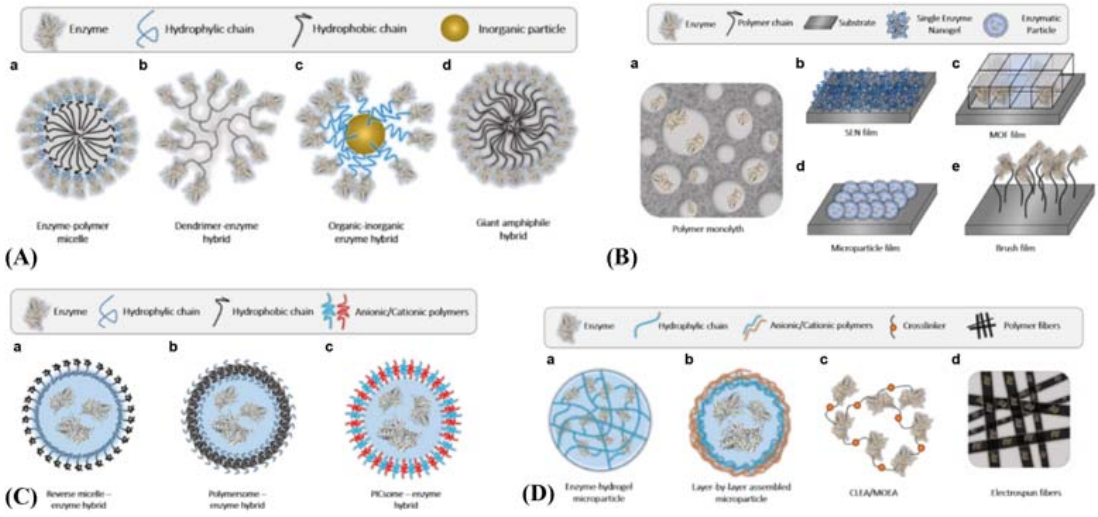


Figure 3.6 Overview on modulation of immobilized enzymes based on polymers for various industrial scale applications. (A) Assembled enzyme polymer supramolecular hybrids in which the enzyme is displayed to the environment. (a) Enzyme–polymer micelles, (b) dendrimer–enzyme hybrids, (c) organic–inorganic enzyme hybrids, (d) giant amphiphile structures. (B) Schematic illustration of enzyme–polymer monoliths and enzyme–polymer film types formed following different approaches mentioned in this work. (a) Polymer monoliths, (b) assembly of Single Enzyme Nanogels (SENs), (c) assembly of Metal Organic Frameworks (MOFs), (d) assembly of microparticles, (e) attachment to polymer brushes. (C) Assembled enzyme–polymer supramolecular hybrids in which the enzyme is confined in the inner cavity of the structures. (a) Reverse micelle–enzyme hybrid, (b) polymersome–enzyme hybrid, (c) PICsome–enzyme hybrid. (D) Schematic illustration of enzyme–polymer–microhybrids described in this work. (a) Enzyme–hydrogel microparticles, (b) microparticles or vesicles formed by layer-by-layer assembly, (c) CLEAs or MOEAs, (d) enzyme polymer fiber hybrids. *CLEAs*, Cross-linked enzyme aggregates; *MOEAs*, metal organic enzyme aggregates. *Adapted from Rodriguez-Abetxuko A, Sánchez-deAlcazar D, Muñumer P, Beloqui A. Tunable polymeric scaffolds for enzyme immobilization. Front Bioeng Biotechnol 2020;8:1–27.*

different types of enzymes. There is continuous updating of immobilization protocols to obtain immobilized enzymes that are easily compatible for various large-scale industrial processes. The prime focus during the development of any new immobilization protocols on percent enzyme recovery, operational stability, selectivity, and reduction in inhibition by products or any other component present in the media. Further, it involves intensive optimization of immobilization solutions to control interaction between matrix and the enzyme and proper orientation of enzyme to maintain its catalytic properties. It is extremely important that there should be continuous stirring of solution containing immobilized enzyme, substrate, and released product at optimum velocity to minimize diffusional constrains by large extent. Experimental details like types of reaction, reaction medium, reaction system, reactor configuration, economic viability, and intrinsic enzyme properties should be thoroughly studied during the designing of robust immobilized enzyme system that can be best suited for any given application [39] (Fig. 3.6). It has been continuously strived to look for the most crucial factor responsible for majority of physicochemical properties of synthesized immobilized enzymes. Till date, the following factors with reported example assured their significance during enzyme immobilization:

1. Immobilization methods and conditions
2. Type of matrix and its properties
3. Nature of enzymes (purity and source of enzyme)
4. Spacer used for binding
5. Type of microenvironment

There are number of reports mentioning significance of various parameters and their optimization during enzyme immobilization. However, it is still known to be method of hit and trial as there is not a single parameter that can predict changes brought up in the enzyme (especially physicochemical properties) during immobilization. It has not yet determined that type of matrix compatibility with bioreactor design and extent of scaling of enzyme-based processes. Sometimes, there are worst outcomes during enzyme immobilization like inactivation of enzyme leading to negligible catalysis despite of enzyme attachment onto/into immobilizing matrices. Thus complete exercise of enzyme immobilization goes vain in addition to wastage of chemicals, time, and labor. Following are few recommendations to avoid complete enzyme inactivation during immobilization:

1. Implementation of computer modeling to predict efficacies of any given immobilizing matrix leading to various enzymatic changes being brought up during enzyme immobilization. It requires 3D structure of enzyme (can be taken from PDB: protein data bank), not available for all enzyme proteins followed by artificial binding of enzyme and matrix through molecular dynamics simulations. However, it still requires wet lab experimentations to confirm results so obtain from dry lab. These studies cannot predict effects of microenvironment on catalysis by immobilized enzymes [40].

2. Molecular impinging has been useful as it helps in controlling conformational selectivity of immobilized enzymes through ligand freezing using physical or chemical methods (namely, lyophilization, cross-linking, and molecular confinement) [41].
3. It has been recommended to use thermophilic enzymes (irrespective of source) that have natural stability during enzyme immobilization, in order to have optimization focus on enhancement of catalytic efficiency.
4. Enzyme stabilization by immobilization through conformational rigidification involving introduction of extra bonds by
   a. chemical modification of enzymes' side chain,
   b. increasing number of hydrogen bonding between enzyme and the matrix, and
   c. manipulation of enzymes' microenvironment.

However, increased stability should not be compensated with any loss in catalytic properties.

1. Orderly arrangement of enzyme molecules onto immobilizing matrix has been useful in providing better stability and activity with respect to randomly immobilized enzymes as it protects the catalytic site of enzyme from inactivation [42].
2. Conformational flexibility (important for catalytic efficiency of immobilized enzymes) should be enhanced by appropriate management of water to organic solvent ratio.
3. There should be appropriate length of spacer (used to attach enzyme to matrix) to prevent significant loss in enzyme activity.
4. There should not be very large number of bonds between enzyme and the matrix, as higher the number of bonds lower would be the enzyme activity (however, greater retentive capability) [43].

3.3.1 Polymeric micelles

1. Size varies from 100 to 200 nm [44].
2. They have central hydrophobic core (made of amphiphilic block copolymers) and hydrophilic outward shell.
3. They are used for enzyme immobilization, specifically attached toward outer shell taking shape of corona.
4. Most of the cases reporting immobilized enzymes have found improved physicochemical (namely, thermostability, operational stability, storage stability, and reusability) and biological (namely, catalytic turnover, reaction velocity, and substrate affinity) properties by providing appropriate environment for the enzymes.
5. Few successful examples are as follows:
   a. Polymeric micelles of 100 nm have been prepared with hydrophilic outer shell made of POEGMA {[poly(oligoethylene glycol)] methyl ether methacrylate}

and central hydrophobic core made of pTBA [poly(*t*-butyl acrylate)]. They are used for the immobilization of industrially important recombinant enzymes (namely, horse radish peroxidase and esterase), especially with his-tag at their C-termini- [45].

b. The Pluronic F127 (polypropylene oxide)-based polymeric micelles have been used for the immobilization of organophosphate hydrolase (OPH) enzyme leading to its improved activity and stability (under conditions like room- and high temperature, freezing and thawing, lyophilization, the presence of organic solvents, etc.) with respect to soluble enzyme [46].

3.3.2 Dendrimers

1. Their size ranging from 100 to 200 nm [47].
2. They are composed of multiple branched polymers emanating radially from a central hydrophobic core.
3. Enzymes are immobilized onto the multiple branches (with high density of functional end groups) of dendrimers through covalent bonding.
4. Few successful examples:
   a. Dendrimers like polyester and poly(amidoamine) (PAMAM) are used to immobilize enzymes [namely, glucose oxidase (GOx) and lipase]. It has been reported that enzyme immobilized onto PAMAM-based dendrimer can be used for delivery purpose in biomedicine [48].
   b. The enzyme immobilization onto dendrimer PBA (phenylboronic acid) through complexation between cationic amine and imidazole groups present on enzymes and boronate present on dendrimer. PBA has been very useful in efficient cytosolic delivery of proteins (namely, Cas9/sgRNA, trypsin, β-galactosidase, lysozyme, cytochrome C, horseradish peroxidase, and RNAse A) with different molecular mass and isoelectric points [49].
   c. PAMAM dendrimer scaffold functionalized with DPA (dipicolylamine)/zinc (II) complex is used for enzyme delivery with great efficiency. It has been commercialized and known by names like TransEx and PULSin delivery systems [50].

3.3.3 Nanoparticles

1. Inorganic nanomaterials (size: 50−200 nm) with a capping layer composed of organic polymers (namely, acrylamide, cellulose, and chitosan) are used for enzyme immobilization on their surface via covalent or noncovalent methods [51].
2. Polymeric coating helps in providing favorable environment to the enzymes and, thus helps in enhancing catalytic activity and recyclability.

3. Generally, inorganic nanomaterials are magnetic in nature which is composed of magnetite and maghemite. They have low toxicity with high enzyme loading capacity and provide easier separation from reaction mixture.
4. It has been found that the nature and the length of polymeric coats as well as size of nanoparticles followed by immobilization methodology have significant impact on the catalytic activities of immobilized enzymes. Further, it has been reported that short and rigid polymeric coating are found to reduce enzymatic activities of immobilized enzymes due to limited flexibility and dynamics.
5. Few examples:
   a. The magnetic nanoparticles with polymeric capping made of poly(GMA) has been used for the immobilization of alcohol dehydrogenase by using glutaraldehyde (cross-linker) and labeled as PGMA-MNPs-ADH. The immobilization efficiency of 97% has been achieved with 99% of enantiomeric excess within 60 min. Immobilized enzyme has reusability of 14 cycles with retention of 80% of initial activity. Same matrix has been used for various other enzymes like *O*-acetylserine sulfhydrylase having his-tag at its *N*-terminus [52].
   b. The nonmagnetic nanoparticles like metals (Zn, Ag, Au, etc.) or silica have been useful with polymeric coating for various enzymes. The enzymes like *C. antarctica* lipase B, *C. antarctica* lipase A and OPH (organophosphate hydrolase) are covalently linked to polymeric shell of metal mesoporous nanoparticles. This immobilization system has been used for the degradation of organophosphate nerve agent (methyl parathion) in aqueous solution as well as 1-phenylethylamine and β-amino ester (ethyl 3-amino-3-phenylpropanoate) in organic solvents with excellent efficiencies (∼98%) [53].

3.3.4 Giant amphiphiles

1. They are used for enzyme immobilization with applications in sensing and heterogeneous catalysis [54].
2. Here, enzyme−polymer conjugates undergo self-assembly leading to giant amphiphilic spherical structures with enzymes at their head (polar) and synthetic polymers at their tail (nonpolar). They resemble like giant surfactants with protein (enzyme) covering interacting with surrounding aqueous solutions and have hydrophobic inner core.
3. The most interesting aspect of these structures is they allow easier control over their morphology and enzyme orientation. Therefore parameters like operational stability, storage stability, and catalytic activity of immobilized enzymes can be improved as per applications.
4. Recently, enzymes like GOx and horseradish peroxidase are collectively taken to prepare enzyme-PHPMA {poly[*N*-(2-hydroxypropyl) methacrylate]} conjugates

followed by their self-assembly through PICA (polymerization-induced coassembly) approach. These giant comicelles have 4.9-fold cascade activity enhancement with respect to soluble-free enzymes. They are much effective for efficient and faster glucose detection [55].

3.3.5 Reverse micelles

1. These reverse polymer micelles have spherical structure (size 10–50 nm) with hydrophilic heads oriented toward core (due to dipole–dipole and ion–dipole interactions), while tails are hydrophobic in nature [56].
2. During enzyme immobilization, a small amount of water is added near to polar heads which helps in enzyme attachment onto these reverse micelles. Usually immobilized enzymes based on these matrices carry out reactions with hydrophobic substrates (namely, lipase, horse radish peroxidase, and α-chymotrypsin). Further, hydrophobic polymer outer shell provides stability to micelle as well as prevents enzymes from denaturation.
3. Various polymers being used for the synthesis of reverse micelles are CTAB (cationic cetyltrimethyl-ammonium bromide), AOT [anionic bis(2-ethylhexyl) sulfosuccinate], Tween-85 (nonionic polyoxyethylene sorbitan trioleate), etc. [57].
4. Immobilized enzymes have shown enhanced catalytic properties with respect to soluble enzymes due to least distortion to conformations of enzymes, lowest mass transfer limitations, and intact active site.

3.3.6 Polymersomes

1. Their size ranges from 100 to 400 nm, having vesicle-like structures (mimic liposomes) composed of amphiphilic block copolymers [58].
2. Enzyme immobilization involves encapsulation inside these polymersomes. Immobilized enzymes are well protected inside polymeric structures from harsh environment and confine their reaction within a nanospace. Thus they provide a platform for multienzyme immobilization with each enzyme being confined to particular compartment.
3. Most of the reports have emphasized on using porous as well as stimuli-responsive membranes for the synthesis of polymersomes for enzyme immobilization to minimize diffusional constrains.
4. Most commonly used polymers for the synthesis of polymersomes are PMOXA [poly(2-methyl-2-oxazoline)], PEG [poly(ethylene glycol)], PIAT {poly[isocyano-L-alanine (2-thiophen-3-ylethyl)amide]}, polydimethylsiloxane, PS (polystyrene), etc. [59].
5. The reported enzymes being immobilized into these polymersomes are peroxidases, oxidases, dehydrogenases, catalases, etc.

3.3.7 PICsomes (polyion complex vesicles)

1. They are based on nanometric polymeric capsules, held together by noncovalent interactions [60].
2. Enzymes are immobilized onto them through electrostatic interactions.
3. Semipermeable membrane–like structure outside these PICsomes helps in easier mass transfer (uptake of substrates, release of products); thus immobilized enzymes have least perturbations toward its catalytic properties with respect to soluble enzymes.
4. The enzymes like asparaginase and β-galactosidase immobilized inside PICsomes are being commercially used for delivery applications [61].

3.3.8 MOF-nanohybrids

1. Size ranges from 500 nm to 2 mm [62].
2. They are porous network of nanoscale coordination polymers (N), well accepted for enzyme immobilization.
3. Their synthesis involves conditions like the presence of organic solvents and high temperatures. Enzymes are immobilized onto surface of these structures through physical adsorption or covalent binding. It is extremely important to determine accurate size of these structures to be implemented for the immobilization of any enzyme to avoid limitation due to pore size.
4. Immobilized enzymes have been used for wide applications like sensing, microfluidic systems, bioremediation, photodynamic therapies, protein delivery or even for the optimization of the polymerase chain reaction.
5. Sometimes divalent metals (Zn^{2+}, Fe^{2+}, Cu^{2+}, etc.) are incorporated into these structures helpful for the immobilization of enzymes like catalase, lipase, horse radish peroxidase, cytochrome *c*, β-galactosidase, and caspase-3 [63].

3.3.9 Conductive polymers

1. They have been very useful for the development of biosensors (namely, detection of biological molecules like glutamic acid, acetylcholine, and soluble gasses as NO associated with various diseases) as they have excellent electrical conductivity, low ionization potential, high electronic affinities, and optical properties [64].
2. Commonly used conductive polymers are polyacetylene, polyaniline, polypyrrole, polythiophene, poly(*p*-phenylene), poly(phenylene vinylene), etc.
3. Few examples:
   a. The enzymes acetylcholinesterase (AChE) and choline oxidase are immobilized onto film {composed of Fe_2O_3NPs–PEDOT [poly(3,4-ethylenedioxythiophene)] and rGO (reduced GO)} followed by its deposition onto fluorine-doped tin oxide and used for the detection of acetyl choline (as low as 4 nM). Thus it is

helpful for Alzheimer's patients (acetylcholine concentration: 5.0−7.8 nM) and discriminating from healthy individuals (acetylcholine concentration: 8.2−11.3 nM). This biosensor has stability of over 60 days with retention of 91% of initial enzyme activity [65].

b. The enzyme GOx has been immobilized onto PtNPs-PEDOT [platinum nanoparticles-poly(3,4-ethylenedioxythiophene)] and used as high sensitive glucose biosensors, useful for diabetic patients [66].

3.3.10 Polymeric brushes

1. These are generally used in modifying surface of the materials, that is, electrodes or particles through deposition [67].
2. These are long and flexible polymer chains with excellent grafting density. Further, parameters (important for enzyme immobilization) like surface chemistry, molecular weight, and grafting density can easily be modified according to applications.
3. These polymeric brushes help in improvising surface area through tunable density of various functional groups like epoxide, carboxylic acid, hydroxyl, aldehyde, and amine groups, used for enzymes' conjugation.
4. Several reports have found that hydrophilic and zwitterionic polymeric brushes are best for enzyme immobilization as they provide ideal microenvironment [68].
5. The number of reactive groups present on polymer brushes should be well modulated as their high density would lead to reduced degree of conformational freedom of the enzyme leading to poor catalytic activities.

3.4 Accorded states of immobilized enzyme onto polymers for industrial applications

3.4.1 Microstructures

The immobilization of enzymes onto/into microscopic polymeric structures leading to their improved catalytic properties with least diffusional constrains has fairly good upscaling capacity (useful in designing biocatalytic bioreactors). They are easily compatible for various industrial processes and proved to have contributed substantially in lowering production cost by several folds. Following sections have provided an overview on enzyme immobilization onto these structures and their applications [69].

3.4.1.1 Enzyme hydrogel microhybrids

1. Enzyme immobilization onto polymeric hydrogels (3D network structures) which can be either natural or synthetic polymers [70].
2. Immobilizing matrix so chosen has tendency to soak large quantity of water and swell leading to exposure of functional hydrophilic groups like $-OH$, $-CONH_2$,

and SO_3H being used for enzyme immobilization via covalent and noncovalent bonds.

3. Commonly used natural polymers for enzyme immobilization: alginate, dextran, gelatin, chitosan, agarose (nontoxic, biocompatible, biodegradable, and provide numerous reactive sites for enzyme attachment).
4. Commonly used synthetic polymers for enzyme immobilization: PVA [poly(vinyl alcohol)], PEI [poly(ethylene imine)], PVP (polyvinylpyrrolidone), PEG, and PNIPAAM [poly(*N*-isopropylacrylamide)]. They are useful as they can be modified to much extent to modulate enzyme activity in its most active state.
5. Most of the reports based on enzyme immobilized onto hydrogel microparticles are being used for applications like biosensors, and enzyme-driven micromotors.
6. Here architecture of the immobilizing matrix has been useful for enzymes by imparting excellent operational and storage stabilities and have high enzyme loading capacities. Further, there are least diffusional strains leading to improve reaction kinetics by immobilized enzymes.
7. Immobilized enzymes have size ranging from 10 to 100 nm.
8. Some reports have used the term click chemistry involving sequential thiol-ene and bioorthogonal tetrazine−norbornene reactions (commonly used for enzyme attachment onto hydrogel-based matrices).

3.4.1.2 Layer-by-layer-wise assembled enzyme−polymer microhybrids

1. Here oppositely charged polyelectrolytes are assembled layer by layer on solid support (usually charged surface or an inorganic template like calcium carbonate) [71].
2. Enzyme can be conveniently immobilized onto these layer-by-layer assembled polymers without losing any of its activity and attains much higher stability with respect to soluble enzyme.
3. Commonly used polymers are PSS (polystyrene sulfonate), PAH [poly(allylamine hydrochloride)], PEI, alginate, etc.
4. Various enzymes like catalase, GOx, α-chymotrypsin, and BGL have been successfully immobilized onto these assembled polymers.
5. Factors like composition, thickness, charge, and permeability of assembled polymers have great impact on activity of immobilized enzymes.
6. Layer-by-layer assembled polymers can be molded into various designs like microspheres, capsules, or hollow capsules according to various applications.
7. Recently, an innovative technique has been developed involving formation of layer-by-layer polymer assembly onto the surface of enzyme crystals leading to better enzyme hydration after polymer assembly while being trapped inside polymeric assembly capsule [72].

3.4.1.3 Cross-linked enzyme–polymer conjugates

1. Often termed CLEAs [73].
2. It involves covalent linking of enzyme molecules in the presence of cross-linker (e.g., glutaraldehyde) leading to CLEAs with excellent conversion efficiencies, good stabilities, and much cheaper than any other immobilized enzyme systems.
3. It involves amino acid–like residues for cross-linking; therefore enzyme molecules with least number of lysine often results in ineffective CLEAs. Generally, various additives like PEIs, PEI-sulfate dextran, dodecyl aldehyde, bovine serum albumin, egg albumin, polylysine, and soy protein isolate are recommended for enzyme molecules with lesser number of lysines.
4. Various biopolymers like dextran-polyaldehyde, gum Arabic, chitosan, and ethylene glycol-bis (succinic acid *N*-hydroxysuccinimide) have been introduced into CLEAs to elongate further operational stabilities and increase a number of reusability cycles to make CLEAs commercially more acceptable. It also helps in reducing high toxicity and reactivity caused to the presence of glutaraldehyde (used for cross-linking) [74].
5. Several innovations have been done to enhance the efficacy of CLEAs, for example, synthesis of MOEAs (metal organic enzyme aggregates) by cross-linking enzyme nanogels via metalimidazole coordination. It has least diffusional constrains and allow higher enzyme loadings, useful for various commercial applications [75].

3.4.1.4 Enzyme–polymer fiber hybrids

1. Polymeric fibers like PVA, PVA/PA6 (PVA/polyamide 6), PAN (polyacrylonitrile), AOPAN (amidoxime PAN), and PEDOT/PAN [poly(3,4-ethylenedioxythiophene)/PAN] have been used to produce nanofibers by electrospun technology and used for enzyme immobilization [76,77].
2. These electrospun nanofibers have high degree of porosity, large number of functional groups (rich in groups like $-NH_2$ and $-OH$) and have excellent biocompatibility. These features are responsible for very good immobilization efficiencies of various enzymes being immobilized onto them by entrapment or surface attachment (adsorption or covalent binding).
3. Immobilized enzymes based on polymeric fibers are used for various applications such as biosensors, pollution control, and bioreactors.
4. Example: enzyme laccase has been immobilized onto PLCL [poly(L-lactic acid)-*co*-poly(+-caprolactone)] electrospun nanofibers by entrapment and used for biodegradation of hazardous pollutants from wastewater. Immobilized laccase has higher affinity for the substrate with respect to soluble enzyme, thus much effective in the removal of pollutants. It has retained 90% of initial activity after 40 days of storage at room temperature [78].

3.4.2 Macrostructures

Enzyme immobilization onto polymers assembled into continuous films, monoliths, etc. resulting into macrostructures with excellent mechanical properties and can be easily modulated according to various applications. Following sections have provided an overview on immobilized enzymes onto/into these structures.

3.4.2.1 Polymer monoliths–enzyme hybrids

1. Enzyme immobilization onto polymer monoliths has been developed in 1990, based on matrices like methacrylates, acrylates, and styrenes [79].
2. Here enzyme immobilization is being performed onto templates like rod polymer for chromatographic columns, disks, and microfluidic channels.
3. Enzymes are usually attached onto polymeric monoliths through covalent bonding based on various functional groups like epoxy, aldehyde, or alkene groups.
4. Immobilized enzymes have excellent operational and storage stabilities in aqueous as well as organic media at room temperature.
5. Enzymes immobilized onto polymeric monoliths are commonly used in biosensors, HPLC (high-performance liquid chromatography), proteomics [namely, Electrospray Ionization Mass Spectrometry (ESI-MS)], microfluidic applications, bioreactors, etc.
6. Example: enzyme trypsin has been immobilized onto polymeric monolith fabricated with ethylene glycol dimethylacrylate (EDMA) based on thiol-ene click chemistry. Immobilized trypsin has K_m of 2.1 mM (soluble enzyme has K_m of 1.4 mM) with reusability cycles of five with retention of 80% initial activity. Trypsin immobilized onto trimethylolpropane trimethacrylate is found to have better catalytic activity and stability. Enzymes like arginase and AChE immobilized onto EDMA are useful for diagnosis of cardiovascular diseases and Alzheimer's, respectively [80].
7. However, polymeric monoliths used for enzyme immobilization often allow lower enzyme loading due to their large pore size. It has been circumvented to much by grafting a polymeric layer onto macropores of monoliths.

3.4.2.2 Enzyme–polymer films

1. These are most commonly used for biosensors, biofuel cells, multienzymatic cascade reactions, etc. [81].
2. Here enzymes are attached onto polymeric films usually by simple deposition, drop casting, spin coating, layer by layer, electro-spinning, dip coating, Langmuir–Blodget technique, etc.
3. Sometimes divalents like Cu^{2+} and Mn^{2+} can be used for imparting extra stability to immobilizing enzyme. For example, enzyme GOx is immobilized in an orderly

manner onto polymeric films using Ca^{2+}, enzyme tyrosinase has been immobilized onto polymeric films in the presence of Cu^{2+}, and used in bisphenol A biosensor, etc. [82].

3.5 Conclusion

Enzyme-based catalysis has been industrially very useful. Industries like pharmaceuticals, food and beverage, biofuel production and natural gas conversion have significant impact on the economy of any nation. Further enzyme-based process can be performed under milder process conditions, least waste generation with least toxicity and have high product sensitivity. It has always been desired to have higher turnover with least cost during production. Enzyme immobilization has been a boon in this respect by imparting operational and storage stabilities to immobilized enzymes with increased number of reusability cycles leading to effective process efficiency. However, enzyme immobilization may lead to the loss of catalytic efficiency due to diffusional constrains and inactivation of active site. There is continuous improvisation in field of enzyme immobilization to minimize mentioned problems. It has been found that immobilizing matrices have huge impact on catalytic efficiency of immobilized enzymes. Some groups have recommended in following protein engineering to bring changes that would allow help in preparing enzyme with additional stability and higher catalytic efficiency with respect to native enzyme. However, it has not been as successful as predicted due to various reasons. Few of them are tedious and do not provide a complete assurance of obtaining super enzyme with mentioned additional features, require lots of expertise, less time, a constant number of reusable enzymes and less cost. Therefore it has never come to commercial market for various industrial processes. Intensive studies on the properties of immobilizing matrices have been carried out widely to improvise them in order to obtain immobilized enzymes with excellent physicochemical properties. The primary focus of commercial scale processes to adapt developed process with compatible economics leading to higher annual turnover. Future prospects of enzyme immobilization has anticipated that structural details of enzymes would be taken and carry out dry experiments to work on compatibility of enzymes and immobilizing matrices followed by expected change in physicochemical properties. In addition, obtained results would be corroborated with wet lab experiments with smaller number of runs leading to most efficient immobilized enzyme systems.

References

[1] Dwevedi A. Enzyme immobilization: an important link between agriculture and industries. In: Enzyme immobilization: advancement in industry, agriculture, medicine and environment. New York: Springer Nature; 2016. p. 45–64. ISBN: 978-3-319-41416-4.

[2] Dwevedi A. 100 Years of enzyme immobilization. In: Enzyme immobilization: advancement in industry, agriculture, medicine and environment. New York: Springer Nature; 2016. p. 1–20. ISBN: 978-3-319-41416-4.
[3] Dwevedi A. Assessment of combo-technology based on its sustainability and life cycle implications. In: Nanotechnology and enzyme technology, combined to solve environmental problems. Elsevier; 2019. p. 151–78. ISBN: 978-0-12-813123-7.
[4] Dwevedi A. Overview of combo-technology (nanotechnology and enzyme technology) and its updates. In: Nanotechnology and enzyme technology, combined to solve environmental problems. Elsevier; 2019. p. 1–47. ISBN: 978-0-12-813123-7.
[5] Dwevedi A. Basics of enzyme immobilization. In: Enzyme immobilization: advancement in industry, agriculture, medicine and environment. New York: Springer Nature; 2016. p. 21–44. ISBN: 978-3-319-41416-4.
[6] Dwevedi A, Kayastha AM. Enzyme immobilization: a breakthrough in enzyme technology and boon to enzyme based industries. In: Haggerty LM, editor. Protein science and engineering. New York: Nova Publishers; 2012. p. 389–419. ISBN: 978-1-61209-656-8.
[7] Singh DP, Dwevedi A. Production of clean energy by green ways. In: Nanotechnology and enzyme technology, combined to solve environmental problems. Elsevier; 2019. p. 49–90. ISBN: 978-0-12-813123-7.
[8] Meghwanshi GK, Kaur N, Verma S, Dabi NK, Vashishtha A, Charan PD, et al. Enzymes for pharmaceutical and therapeutic applications. Appl Biochem Biotechnol 2020;67:586–601.
[9] Alcantara CM, Alcantara AR. Biocatalyzed synthesis of antidiabetic drugs: a review. Biocatal Biotransform 2017;36. Available from: https://doi.org/10.1080/10242422.2017.1323887.
[10] de María PD, de Gonzalo G, Alcántara AR. Biocatalysis as useful tool in asymmetric synthesis: an assessment of recently granted patents (2014–2019). Catalysts 2019;9:1–42.
[11] Howe AYM, Venkatraman S. The discovery and development of boceprevir: a novel, first-generation inhibitor of the hepatitis C virus NS3/4A serine protease. J Clin Transl Hepatol 2013;1:22–32.
[12] Wienkers LC, Allievi C, Hauer MJ, Wynalda MA. Cytochrome P-450-mediated metabolism of the individual enantiomers of the antidepressant agent reboxetine in human liver microsomes. Drug Metab Dispos 1999;27:1334–40.
[13] Winkler CK, Clay D, Davies S, O'Neill P, McDaid P, Debarge S, et al. Chemoenzymatic asymmetric synthesis of pregabalin precursors *via* asymmetric bioreduction of β-cyanoacrylate esters using ene-reductases. J Org Chem 2013;78:1525–33.
[14] Jegannathan KR, Abang S, Poncelet D, Chan ES, Ravindra P. Production of biodiesel using immobilized lipase—a critical review. Crit Rev Biotechnol 2008;28:253–64.
[15] Sánchez-Ramírez J, Martínez-Hernández JL, Segura-Ceniceros P, López G, Saade H, Medina-Morales MA, et al. Cellulases immobilization on chitosan-coated magnetic nanoparticles: application for *Agave atrovirens* lignocellulosic biomass hydrolysis. Bioprocess Biosyst Eng 2017;40:9–22.
[16] Khoshnevisan K, Poorakbar E, Baharifar H, Barkhi M. Recent advances of cellulase immobilization onto magnetic nanoparticles: an update review. Magnetochemistry 2019;5:1–22.
[17] Han J, Wang L, Wang Y, Dong J, Tang X, Ni L, et al. Preparation and characterization of Fe_3O_4-NH_2@4-arm-PEG-NH_2, a novel magnetic four-arm polymer-nanoparticle composite for cellulase immobilization. Biochem Eng J 2018;130:90–8.
[18] Sheldon RA. CLEAs, Combi-CLEAs and 'smart' magnetic CLEAs: biocatalysis in a bio-based economy. Catalysts 2019;9:1–30.
[19] Song Q, Mao Y, Wilkins M, Segato F, Prade R. Cellulase immobilization on superparamagnetic nanoparticles for reuse in cellulosic biomass conversion. AIMS Bioeng 2016;3:264–76.
[20] Gao J, Lu C-L, Wang Y, Wang S-S, Shen J-J, Zhang J-X, et al. Rapid immobilization of cellulase onto graphene oxide with a hydrophobic spacer. Catalysts 2018;8:1–12.
[21] Ahmad R, Sardar M. Immobilization of cellulase on TiO_2 nanoparticles by physical and covalent methods: a comparative study. Indian J Biochem Biophys 2014;51:314–20.
[22] Cabrera MP, Assis CRD, Neri DFM, Pereira CF, Soria F, Carvalho LB. High sucrolytic activity by invertase immobilized onto magnetic diatomaceous earth nanoparticles. Biotechnol Rep 2017;14:38–46.

[23] Bayramoglu G, Doz T, Ozalp VC, Arica MY. Improvement stability and performance of invertase *via* immobilization on to silanized and polymer brush grafted magnetic nanoparticles. Food Chem 2017;221:1442–50.
[24] Waifalkar PP, Parit SB, Chougale AD, Sahoo SC, Patil PS, Patil PB. Immobilization of invertase on chitosan coated γ-Fe_2O_3 magnetic nanoparticles to facilitate magnetic separation. J Colloid Interface Sci 2016;482:159–64.
[25] Valerio SG, Alves JS, Klein MP, Rodrigues RC, Hertz PF. High operational stability of invertase from *Saccharomyces cerevisiae* immobilized on chitosan nanoparticles. Carbohydr Polym 2013;92:462–8.
[26] Agrawal R, Srivastava A, Verma AK. Immobilization of β-glucosidase onto silicon oxide nanoparticles and augment of phenolics in sugarcane juice. J Food Sci Technol 2016;53:3002–12.
[27] Chen T, Yang W, Guo Y, Yuan R, Xu L, Yan Y. Enhancing catalytic performance of β-glucosidase *via* immobilization on metal ions chelated magnetic nanoparticles. Enzyme Microb Technol 2014;63:50–7.
[28] Zhou Y, Yuan S, Liu Q, Yan D, Wang Y, Gao L, et al. Synchronized purification and immobilization of his-tagged β-glucosidase *via* Fe_3O_4/PMG core/shell magnetic nanoparticles. Sci Rep 2017;7:1–11.
[29] Verma ML, Chaudhary R, Tsuzuki T, Barrow CJ, Puri M. Immobilization of β-glucosidase on a magnetic nanoparticle improves thermostability: application in cellobiose hydrolysis. Bioresour Technol 2013;135:2–6.
[30] Singh RK, Zhang Y-W, Nguyen N-P-T, Jeya M, Lee J-K. Covalent immobilization of β-1,4-glucosidase from *Agaricus arvensis* onto functionalized silicon oxide nanoparticles. Appl Microbiol Biotechnol 2011;89:337–44.
[31] Gao Y, Kyratzis I. Covalent immobilization of proteins on carbon nanotubes using the cross-linker 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide—a critical assessment. Bioconjugate Chem 2008;19:1945–50.
[32] Dhiman SS, Kalyani D, Jagtap SS, Haw J-R, Kang YC, Lee J-K. Characterization of a novel xylanase from *Armillaria gemina* and its immobilization onto SiO_2 nanoparticles. Appl Microbiol Biotechnol 2013;97:1081–91.
[33] Landarani-Isfahani A, Taheri-Kafrani A, Amini M, Mirkhani V, Moghadam M, Soozanipour A, et al. Xylanase immobilized on novel multifunctional hyperbranched polyglycerol-grafted magnetic nanoparticles: an efficient and robust biocatalyst. Langmuir 2015;31:9219–27.
[34] Saurabh SD, Sujit SJ, Marimuthu J, Jung R, Yun C, Jung L. Immobilization of *Pholiota adiposa* xylanase onto SiO_2 nanoparticles and its application for production of xylooligosaccharides. Biotechnol Lett 2012;34:1307–13.
[35] Hwang IY, Lee SH, Choi YS, Park SJ, Na JG, Chang IS, et al. Biocatalytic conversion of methane to methanol as a key step for development of methane-based biorefineries. J Microbiol Biotechnol 2014;24:1597–605.
[36] Majumder AB, Singh B, Dutta D, Sadhukhan S, Munishwar N, Gupta MN. Lipase catalyzed synthesis of benzyl acetate in solvent-free medium using vinyl acetate as acyl donor. Bioorg Med Chem Lett 2006;16:4041–4.
[37] Badgujar KC, Bhanage BM. Enhanced biocatalytic activity of lipase immobilized on biodegradable copolymer of chitosan and polyvinyl alcohol support for synthesis of propionate ester: kinetic approach. Ind Eng Chem Res 2014;53:18806–15.
[38] Robinson PK. Enzymes: principles and biotechnological applications. Essays Biochem 2015;59:1–41.
[39] Basso A, Serban S. Industrial applications of immobilized enzymes—a review. Mol Catal 2019;479:1–20.
[40] Frushicheva MP, Mills MJL, Schopf P, Singh MK, Warshel A. Computer aided enzyme design and catalytic concepts. Curr Opin Chem Biol 2014;21:56–62.
[41] Hoarau M, Badieyan S, Marsh ENG. Immobilized enzymes: understanding enzyme – surface interactions at the molecular level. Org Biomol Chem 2017;15:9539–51.

[42] Zhang D-H, Yuwen L-X, Peng L-J. Parameters affecting the performance of immobilized enzyme. J Chem 2013;2013:1–7.
[43] Mohamad NR, Marzuki NH, Buang NA, Huyop F, Wahab RA. An overview of technologies for immobilization of enzymes and surface analysis techniques for immobilized enzymes. Biotechnol Biotechnol Equip 2015;29:205–20.
[44] Besic S, Minteer SD. Micellar polymer encapsulation of enzymes. Methods Mol Biol 2017;1504:93–108.
[45] Rodriguez-Abetxuko A, Sánchez-deAlcázar D, Muñumer P, Beloqui A. Tunable polymeric scaffolds for enzyme immobilization. Front Bioeng Biotechnol 2020;8:1–27.
[46] Suthiwangcharoen N, Nagarajan R. Enhancing enzyme stability by construction of polymer–enzyme conjugates micelles for decontamination of organophosphate agents. Biomacromolecules 2014;15:1142–52.
[47] Cardoso FP, Neto SA, Ciancaglini P, de Andrade AR. The use of PAMAM dendrimers as a platform for laccase immobilization: kinetic characterization of the enzyme. Appl Biochem Biotechnol 2012;167:1854–64.
[48] Abedi-Gaballu F, Dehghan G, Ghaffari M, Yekta R, Abbaspour-Ravasjani S, Baradaran B, et al. PAMAM dendrimers as efficient drug and gene delivery nanosystems for cancer therapy. Appl Mater Today 2018;12:177–90.
[49] Liu C, Wan T, Wang H, Zhang S, Ping Y, Cheng Y. A boronic acid–rich dendrimer with robust and unprecedented efficiency for cytosolic protein delivery and CRISPR-Cas9 gene editing. Sci Adv 2019;5:1–11.
[50] Fana M, Gallien J, Srinageshwar B, Dunbar GL, Rossignol J. PAMAM dendrimer nanomolecules utilized as drug delivery systems for potential treatment of glioblastoma: a systematic review. Int J Nanomed 2020;15:2789–808.
[51] Wedegärtner D, Strube OI. Site-specific addressing of particles and coatings *via* enzyme-mediated destabilization. Catalysts 2019;9:1–14.
[52] Bach LG, Islam MR, Lee DC, Lim KT. Poly(glycidyl methacrylate) grafted CdSe quantum dots by surface-initiated atom transfer radical polymerization: novel synthesis, characterization, properties, and cytotoxicity studies. Appl Surf Sci 2013;283:546–53.
[53] Kudr J, Haddad Y, Richtera L, Heger Z, Cernak M, Adam V, et al. Magnetic nanoparticles: from design and synthesis to real world applications. Nanomaterials 2017;7:1–29.
[54] Bao C, Yin Y, Zhang Q. Synthesis and assembly of laccase-polymer giant amphiphiles by self-catalyzed CuAAC click chemistry. Biomacromolecules 2018;19:1539–51.
[55] Arica Y, Hasirci VN. Immobilization of glucose oxidase in poly(2-hydroxyethyl methacrylate) membranes. Biomaterials 1987;8:489–95.
[56] Yi S, Dai F, Zhao C, Si Y. A reverse micelle strategy for fabricating magnetic lipase-immobilized nanoparticles with robust enzymatic activity. Sci Rep 2017;7:1–19.
[57] Jones M-C, Leroux J-C. Reverse micelles from amphiphilic branched polymers. Soft Matter 2010;6:5850–9.
[58] Dinu MV, Dinu IA, Saxer SS, Meier W, Pieles U, Bruns N. Stabilizing enzymes within polymersomes by coencapsulation of trehalose. Biomacromolecules 2021;22:134–45.
[59] Rideau E, Dimova R, Schwille P, Wurm FR, Landfester K. Liposomes and polymersomes: a comparative review towards cell mimicking. Chem Soc Rev 2018;47:8572–610.
[60] Tang H, Sakamura Y, Mori T, Katayama Y, Kishimura A. Development of enzyme loaded polyion complex vesicle (PICsome): thermal stability of enzyme in PICsome compartment and effect of coencapsulation of dextran on enzyme activity. Macromol Biosci 2017;17. Available from: https://doi.org/10.1002/mabi.201600542.
[61] Anraku Y, Kishimura A, Kamiya M, Tanaka S, Nomoto T, Toh K, et al. Systemically injectable enzyme-loaded polyion complex vesicles as *in vivo* nanoreactors functioning in tumors. Angew Chem Int Ed Engl 2016;55:560–5.
[62] Rodriguez-Abetxuko A, Morant-Miñana MC, Knez M, Beloqui A. Carrierless immobilization route for highly robust metal–organic hybrid enzymes. ACS Omega 2019;4:5172–9.

[63] Ahmad R, Shanahan J, Rizaldo S, Kissel DS, Stone KL. Co-immobilization of an enzyme system on a metal-organic framework to produce a more effective biocatalyst. Catalysts 2020;10:1–14.
[64] Lakard B. Electrochemical biosensors based on conducting polymers: a review. Appl Sci 2020;10:1–24.
[65] Tunç AT, Koyuncu EA, Arslan F. Development of an acetylcholinesterase-choline oxidase based biosensor for acetylcholine determination. Artif Cells Nanomed Biotechnol 2016;44:1659–64.
[66] Liu Y, Turner APF, Zhao M, Mak WC. Processable enzyme-hybrid conductive polymer composites for electrochemical biosensing. Biosens Bioelectron 2018;100:374–81.
[67] Koenig M, König U, Eichhorn K-J, Müller M, Stamm M, Uhlmann P. In-situ-investigation of enzyme immobilization on polymer brushes. Front Chem 2019;7:1–8.
[68] Jiang H, Xu F-J. Biomolecule-functionalized polymer brushes. Chem Soc Rev 2013;42:3394–426.
[69] Feng Y, Zhong L, Bilal M, Tan Z, Hou Y, Jia S, et al. Enzymes@ZIF-8 nanocomposites with protection nanocoating: stability and acid-resistant evaluation. Polymers 2019;11:1–14.
[70] Onbas R, Yesil-Celiktas O. Synthesis of alginate-silica hybrid hydrogel for biocatalytic conversion by β-glucosidase in microreactor. Eng Life Sci 2019;19:37–46.
[71] Sakr OS, Borchard G. Encapsulation of enzymes in layer-by-layer (LbL) structures: latest advances and applications. Biomacromolecules 2013;14:2117–35.
[72] Ariga K, Hill JP, Ji Q. Layer-by-layer assembly as a versatile bottom-up nanofabrication technique for exploratory research and realistic application. Phys Chem Chem Phys 2007;9:2319–40.
[73] Cheng H, Zhao Y-L, Luo X-J, Xu D-S, Cao X, Xu J-H, et al. Cross-linked enzyme-polymer conjugates with excellent stability and detergent-enhanced activity for efficient organophosphate degradation. Bioresour Bioprocess 2018;5:1–9.
[74] Bezerra CS, de Farias Lemos CMG, de Sousa M, Rocha L, Gonçalves B. Enzyme immobilization onto renewable polymeric matrixes: past, present, and future trends. J Appl Polym Sci 2015;42125:1–15.
[75] Beloqui A, Cortajarena AL. Protein-based functional hybrid bionanomaterials by bottom-up approaches. Curr Opin Struct 2020;63:74–81.
[76] Zhu X, He B, Zhao C, Ma Y, Yang W. Separated immobilization of incompatible enzymes on polymer substrate *via* visible light induced living photografting polymerization. Langmuir 2017;33:5577–84.
[77] Wang Z-G, Wan L-S, Liu Z-M, Huang X-J, Xu Z-K. Enzyme immobilization on electrospun polymer nanofibers: an overview. J Mol Catal B Enzym 2009;56:189–95.
[78] Ba S, Arsenault A, Hassani T, Jones JP, Cabana H. Laccase immobilization and insolubilization: from fundamentals to applications for the elimination of emerging contaminants in wastewater treatment. Crit Rev Biotechnol 2013;33:404–18.
[79] Krenkova J, Svec F. Less common applications of monoliths: IV. Recent developments in immobilized enzyme reactors for proteomics and biotechnology. J Sep Sci 2009;32:706–18.
[80] Krenkova J, Lacher NA, Svec F. Highly efficient enzyme reactors containing trypsin and endoproteinase LysC immobilized on porous polymer monolith coupled to MS suitable for analysis of antibodies. Anal Chem 2009;81:2004–12.
[81] Cordeiro AL, Pompe T, Salchert K, Werner C. Enzyme immobilization on reactive polymer films. Methods Mol Biol 2011;751:465–76.
[82] Inroga FAD, Rocha MO, Lavayen V, Arguello J. Development of a tyrosinase-based biosensor for bisphenol A detection using gold leaf–like microstructures. J Solid State Electrochem 2019;23:1659–66.

View publication stats

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

Operation and Control of Three-Phase Solar PV System Connected to Distribution Network

Hanish Mehla
Dept. of Electrical Engineering
Delhi Technological University
Delhi, India
hanishmehla@gmail.com

Manish Kumar
Dept. of Electrical Engineering
Delhi Technological University
Delhi, India
manishnareda01@gmail.com

Manish Yadav
Dept. of Electrical Engineering
Delhi Technological University
Delhi, India
manishyadav1729@gmail.com

Madhusudan Singh
Dept. of Electrical Engineering
Delhi Technological University
Delhi, India
madhusudan@dce.ac.in

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498516

Abstract— **This paper presents a solar PV grid integrated system in a distribution network. The integration of the solar PV system, along with the electrical distribution grid, involves two stages. The boost converter is the first stage, wherein its primary function is to assess the optimum power from a PV array. The boost converter works on the principle of MPPT (Maximum Power Point Tracking) and provides an adaptive DC link voltage as an input to the second stage. The second stage consists of a voltage source converter (VSC), which feeds the maximum power extracted from the PV array to the distribution grid by converting DC voltage from PV array into AC voltage. A d-q control approach has been used to control the VSC. The d-q control approach provides an easy control mechanism along with higher bandwidth as compared to tracking controllers. Also, the d-q controller provides a much simpler design, where only one d-q controller is used as compared to tracking controllers wherein three controllers are used (one for each phase). This paper presents modelling, control, and analysis of the grid-tied three-phase solar PV system under varying environmental conditions on solar PV and variations in voltages of distribution systems. The proposed system possesses low Total Harmonic Distortion (THD), i.e., as per the IEEE-519 standard. The performance of the proposed system has been verified for normal grid conditions and under and over voltages in the distribution system through a simulation study in MATLAB.**

Keywords— MPPT, d-q, Adaptive DC-link Voltage

NOMENCLATURE

SPV:	Solar Photo Voltaic
PV:	Photo Voltaic
DC:	Direct Current
AC:	Alternating Current
MPPT:	Maximum Power Point Tracking
THD:	Total Harmonic Distortion
VSC:	Voltage Source Converter
PLL:	Phase Locked Loop
INC:	Incremental Conductance Method
P&O:	Perturb and Observe
Ipv:	PV Array Current
Vpv:	PV Array Voltage
CPI:	Common Point of Interface

I. INTRODUCTION

Solar Energy is a vital form of energy available to humankind right from its earliest times. Earth has an enormous supply of solar energy, and this energy can be made to produce electricity and thus reduce the load on conventional electrical energy-producing systems.

Nowadays, there has been an increase in demand for SPV systems due to their reduced cost and increased efficiency. The recent developments in power electronics devices and converter topology have increased the scope of SPV integration into the grid. Further, the increasing emphasis of the world to integrate SPV systems with the current electrical grid networks [1]-[2] results in reduced emission and reduced generation costs. Emphasis has been laid on solar energy because of the enormous availability of solar energy in India [3]. SPV provides an environmentally friendly choice for electrical energy production since nearly no greenhouse gases are produced during the operation of SPV systems. Integration of a three-phase SPV system with a distribution electrical grid network also reduces the generation from conventional thermal and nuclear power plants.

In this paper, a Solar PV system is integrated with the electrical distribution grid. Since distributed generations provide various advantages right from reducing electrical losses to load profile flattening [4]-[6] and small to medium scale SPV systems can also be integrated into the distribution network, which is installed on roofs of houses/domestic buildings.

There are two types of solar PV systems- Standalone and Grid interfaced systems. Standalone systems work on energy storage (using conventional batteries), whereas grid interfaced systems are directly tied to the grid, wherein the grid serves as the buffer. The voltage and frequency control problems associated with standalone systems [7] make grid-tied SPV systems more preferred.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:54:03 UTC from IEEE Xplore. Restrictions apply.

In this paper, the proposed system consists of two stages- the first stage being of Boost converter and the second one being of the VSC. In the first stage, the boost converter makes the PV array to operate at MPPT.

MPPT Technique is utilized to assess optimum power from the PV array [8]. Various MPPT Techniques are available in the literature like P&O, INC, Fractional short circuit current method, and Fractional open-circuit voltage method. Among all these techniques, INC is preferred due to its simplicity and high efficiency, even during voltage variations [9].

The second stage in SPV integration involves a VSC. An adaptive DC-Link voltage is provided as an input to the VSC. It has various advantages over the conventional constant DC link voltage. The DC Link voltage is made adaptive in accordance with CPI voltage. The d-q control approach is used to provide gate pulses to control the VSC keeping in mind the various advantages it provides like increased bandwidth, simple design and a single controller used compared to three (one for each phase) used in conventional tracking point controller. The PLL-based technique synchronizes and determines the initial phase angle (wt) to compute the d-q axis reference currents.

The three-phase SPV system connected to a distribution network is simulated in MATLAB/Simulink. The system's performance under varying PV grid conditions is discussed. The THD of the system has been found well under the limits described in IEEE-519 standards (under 5%) for under voltages and overvoltages conditions in the grid network.

II. System Configuration

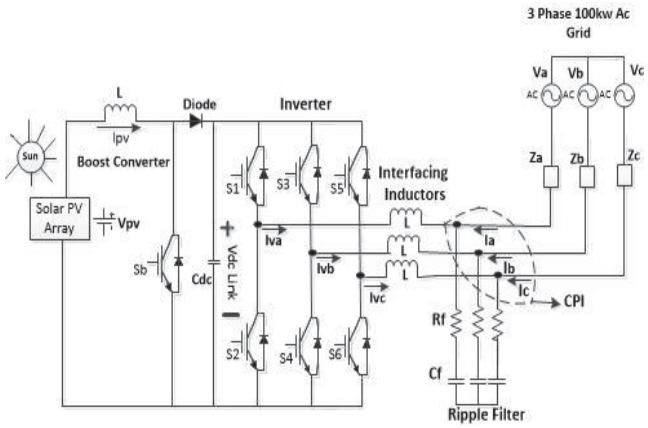


Fig. 1. System configuration

A schematic diagram of the proposed system consisting of SPV Array, Boost converter, Adaptive DC-link Voltage branch, VSC, Interfacing Inductors, Filters and a Distribution Grid is shown in Fig. 1.

The proposed system is consisting two stages- the Boost Converter is the first one, which assists the purpose of MPPT. The input to boost the converter is Ipv and Vpv. The boost converter controls the Vpv in such a way to give maximum power at its output terminals using the INC control approach. The boost converter's output feeds the DC link voltage, which acts as an input to VSC. The DC Link voltage, which is input to VSC, is dynamically adjusted by grid-tied VSC according to CPI voltage. This property overcomes the disadvantages associated with constant DC link voltage.

The second stage is the VSC, and it serves the purpose of converting the DC power to AC and feeding it into the grid. The VSC consists of Three IGBT'S legs with two IGBT'S in each leg and thus act as a three-phase six pulse inverter. The VSC controller handles the control of the VSC. Interfacing Inductors are added to the output of VSC at one end and CPI at the other end to reduce the ohmic losses. The VSC generates high-frequency switching ripples in the output voltages, which can be filtered out by connecting a ripple filter near CPI.

III. Control Approaches

The controlling techniques for boost converter and grid-tied inverter are two primary control approaches in the SPV system. The input voltage of the boost converter is adjusted by implementing the MPPT algorithm to boost converter control. The output of the boost converter, DC-link voltage, is also kept adaptive according to CPI voltage variation. The system is operated so that controlling both the input and output voltage of the boost converter is according to the sensed variables of the circuit. The boost converter feeds power to the DC-link of VSC, which converts the DC power into AC power and inject real and reactive power into the grid. An incremental conductance-based (INC) MPPT algorithm is implemented to estimate the reference PV array voltage. Space vector modulation is used for the control of VSC. Two PI controllers are used as direct and quadrature axis current controllers. Outputs of PI controllers are reference d and q axis voltages converted in reference abc voltages. The reference voltages are then compared with the sawtooth waveform to generated switching pulses for VSC. The basic control approach is shown in Fig. 2. The SPV system's control algorithms are formulated in the followed subsections.

A. Control approach for Boost Converter

An incremental conductance-based MPPT algorithm is used in this paper. The incremental conductance algorithm works on the principle that maximum power is obtained when the difference between conductance offered by the PV array and incremental conductance is minimal. The reference voltage of the PV array is assessed using the incremental conductance principle.

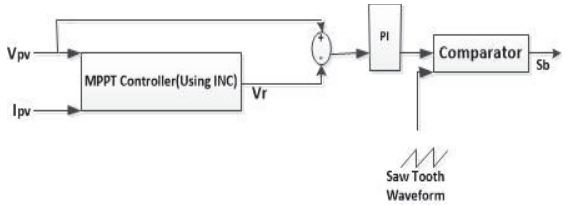


Fig. 2. Control approach of a boost converter

ΔIpv and ΔVpv are estimated for evaluation of incremental conductance as:

$$\Delta Ipv = Ipv(k) - Ipv(k-1) \quad (1a)$$

$$\Delta Vpv = Vpv(k) - Vpv(k-1) \quad (1b)$$

where $Ipv(k)$ and $Vpv(k)$ represent the instantaneous sampled current and voltage of the solar array at k^{th} instant and $Ipv(k-1)$ and $Vpv(k-1)$ are values on the previous instant.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:54:03 UTC from IEEE Xplore. Restrictions apply.

As per the incremental conductance-based MPPT algorithm,

$$\Delta Ipv/\Delta Vpv = -Ipv/Vpv \text{ , at MPP} \quad (2a)$$

$$\Delta Ipv/\Delta Vpv > -Ipv/Vpv \text{ , Left at MPP} \quad (2b)$$

$$\Delta Ipv/\Delta Vpv < -Ipv/Vpv \text{ , Right at MPP} \quad (2c)$$

The (1) and (2) estimates the reference PV array voltage (Vr). The error between reference PV array voltage and measured PV array voltage is given to a PI controller for duty ratio calculation. The duty ratio is the PI controller output, compared with the saw tooth waveform to generate switching logic for the switch in boost converter.

B. Control algorithm for VSC

The primary purpose of VSC control algorithms is to regulate DC-link voltage and feed power to the grid at unity power factor with respect to the CPI. For control of VSC, CPI phase grid voltage (v_a, v_b, v_c) DC link voltage (V_{dc}) and inverter three-phase currents (i_{va}, i_{vb}, i_{vc}) are sensed.

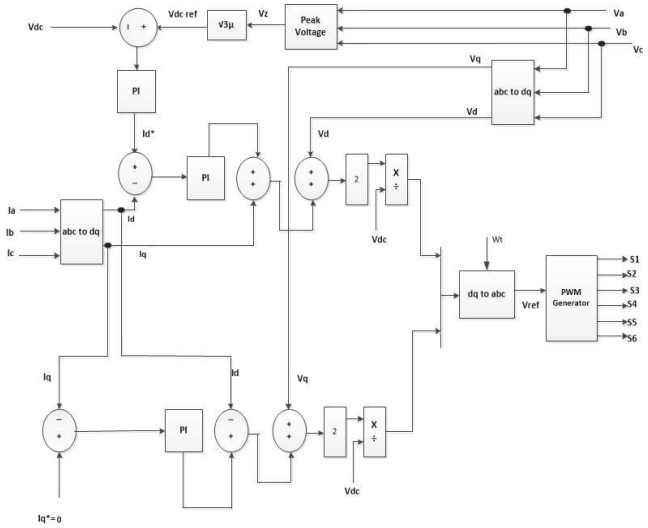


Fig. 3. Control of VSC

The amplitude of AC voltage at CPI is estimated as,

$$V_z = \sqrt{\frac{2(v_a^2 + v_b^2 + v_c^2)}{3}} \quad (3)$$

The VSC output current is appropriately controlled by maintaining the DC-link voltage of VSC higher than the line voltage amplitude to ensure effective control of pulse width modulation (PWM). The reference DC link voltage V_{dcref} is tuned with respect to the CPI voltage amplitude to adjust Vdc with respect to CPI variation adaptively.

The reference DC link voltage V_{dcref} is assessed as:

$$V_{dcref} = \mu\sqrt{3}V_z, where\ \mu > 1 \quad (4)$$

For appropriate current control, the DC-link voltage V_{dc} is set high as to the CPI line voltage amplitude. For that, V_{dcref} is retain approximately 10% high as to the peak of CPI line voltage while the switching drops are taken into account. Hence, the selected value is for μ is 1.1.

The DC link voltage is upheld to V_{dcref} by utilizing a PI controller. The difference between $VdCref$ and Vdc is given to the PI controller. Reference d-axis current i_d^* is obtained at the output of the PI controller.

C. Grid Synchronization

For synchronization of VSC to the grid, the Park transformation is carried out, in which abc reference frame components are transformed into rotating reference or direct-quadrature frame. This three-phase to two-phase conversion is obtained as follow:

$$\begin{bmatrix} d \\ q \end{bmatrix} = \begin{bmatrix} \sin wt & \sin wt - \frac{2\pi}{3} & \sin wt + \frac{2\pi}{3} \\ \cos wt & \cos wt - \frac{2\pi}{3} & \cos wt + \frac{2\pi}{3} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad (5)$$

The Phase Lock Loop (PLL) implementation is the next step for grid synchronization in which the angle (wt) required for the estimation of the above transformation. The block diagram of the PLL is shown in Fig. 4.

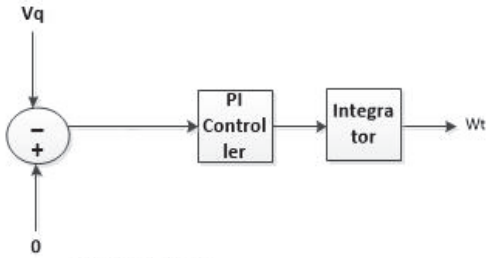


Fig. 4. A block diagram of PLL

As depicted in Fig. 4, the error signal obtained by comparing the quadrature voltage v_q to zero, passed through a PI controller to produce angular frequency. Further, it passed through the integrator to get the instantaneous phase angle (wt).

D. Current Control loop

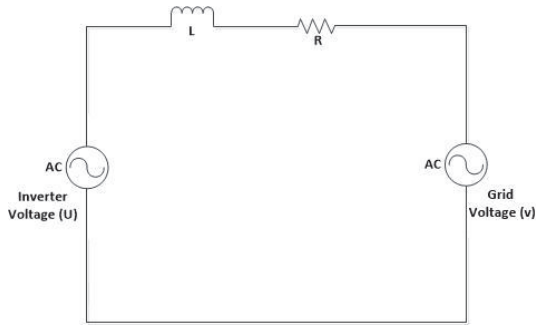


Fig. 5. Current controller

In Fig. 5, the L represents the inductance and a filter, and R represents the total resistance between the inverter and grid. Mathematically, the above circuit can be expressed according to the Kirchhoff voltage law (KVL) as:

$$L\frac{di}{dt} + Ri = \Delta v \quad (6)$$

where $\Delta v = u - v$ and for the three-phase circuit, v and i can be replaced by v_{abc} and $iabc$, respectively. The equations attained after using Park transformation for the conversion of abc to dq and differentiating i to time are:

$$ud = vd + Rid + L\frac{did}{dt} - wLiq \quad (7)$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:54:03 UTC from IEEE Xplore. Restrictions apply.

$$uq = vq + Riq + \frac{diq}{dt} + wLid \quad (8)$$

For independent controlling of P and Q, ud and uq should depend only on d axis components and on q axis components, respectively. The impedance between the grid and inverter causes the voltage drop is compensated by the PI controller, as shown in (7) and (8).

$$ud = vd + (Kp + \frac{Ki}{s})(id* - id) - wLid \quad (9)$$

$$uq = vq + (Kp + \frac{Ki}{s})(iq* - iq) + wLid \quad (10)$$

$iq*$ is assumed zero for considering the unity power factor operation of VSC, where Kp and Ki are gain constants of the proportional and integral term in the PI controller.

To get the modulation indices md and mq, both ud and uq are multiplied with the factor of $(2/V_{dc})$.

$$m_d = ud * \frac{2}{Vdc} \quad (11)$$

$$m_q = uq * \frac{2}{Vdc} \quad (12)$$

m_d and m_q are transformed into abc frame using Park transformation and then given to PWM generators to generate switching pulses for the VSC.

IV. Simulation Results

Simulations are carried out in MATLAB Simulink using sim power system toolbox. The results shown are for the 100kW SPV system. Total harmonic distortion of inverter currents and voltages are well as per IEEE-519 standard (less than 5%) for all varying operating conditions. Simulation results for different operating conditions are shown in Fig. 6.

For the proposed SPV system, the performance is recorded for swift changes in solar insolation from $1000W/m^2$ to $500W/m^2$ as shown in Fig. 7. It can be observed that up to t=0.3 seconds, the solar insolation of the system is $1000W/m^2$ and working under the steady-state condition. Also, the three-phase grid currents i_{abc} are balanced and sinusoidal. After that, at t=0.3 seconds, insolation is swiftly changed from $1000W/m^2$ to $500W/m^2$. The PV array currents I_{pv} and power output P_{pv} decrease due to a decrease in insolation. DC-link voltage V_{dc} shows deviation, but it is quickly settled to its average value. The system quickly settles to the next steady-state and delivers reduced power into the grid.

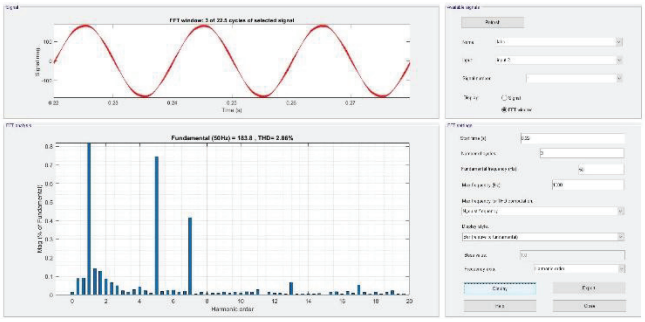


Fig. 6. Total Harmonic Distortion (THD) in injected current into the grid

A. Simulation result for a swift change in solar insolation

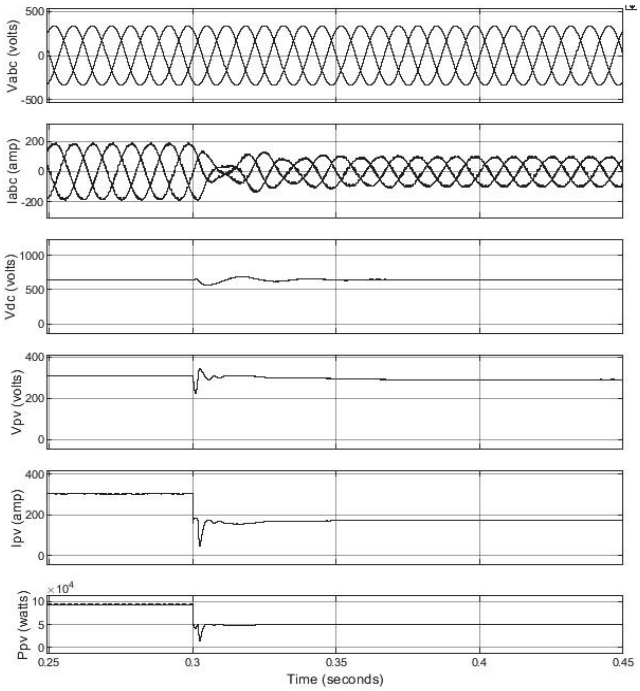


Fig. 7. Performance of the proposed system under a sudden change in solar irradiance

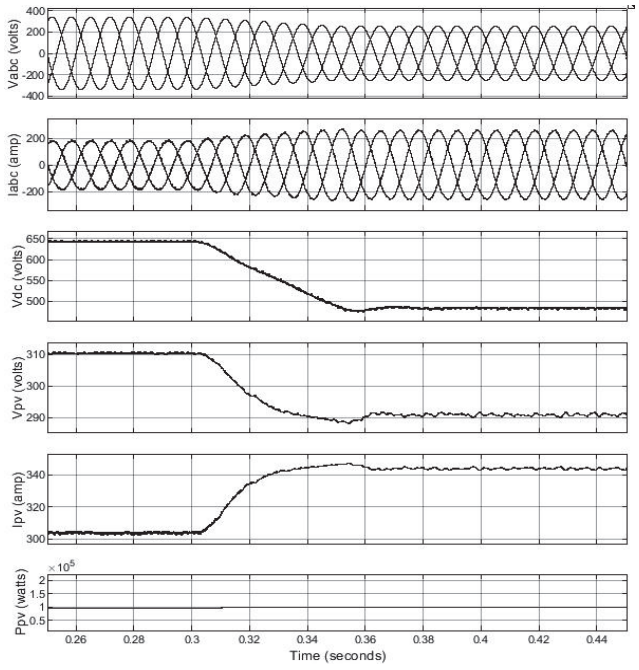


Fig. 8. Steady-state and dynamic performance of the system for under-voltage operation at CPI

B. Simulations result for under-voltage operation

The system's performances are shown in Fig. 8 for under-voltage operation at CPI under the steady-state and dynamic state. It can be observed that up to time t=0.3 seconds, the proposed system is operational at a CPI voltage v_{abc} of 415V. During time 0.3 seconds to 0.35 seconds, CPI voltage is decreased from 415V to 310V. The adaptive nature of the V_{dc} can be noticed in the figure. As CPI voltage decrease results in a decrease in the V_{dc} and to feed

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:54:03 UTC from IEEE Xplore. Restrictions apply.

the matching power to the grid, an increase in inverter currents i_{abc} is observed at reduced voltage.

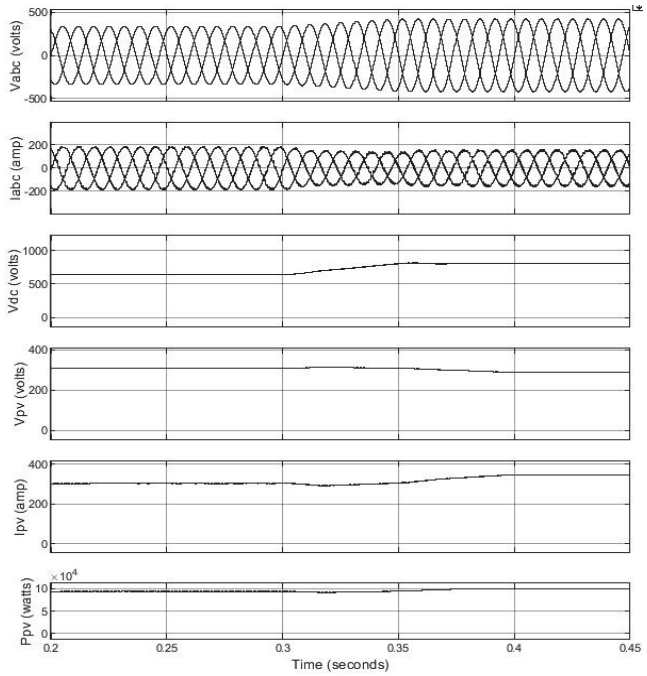


Fig. 9. Steady-state and dynamic performance of the system for over-voltage operation at CPI

C. Simulation results for over-voltage operation

The system's performance under steady-state and dynamic state are shown in Fig. 9 for over-voltage v_{abc} operation at CPI. It can be observed that up to time t=0.3 seconds, the system operates at a CPI voltage of 415V. During time 0.3 seconds to 0.35 seconds, CPI voltage is increased from 415V to 520V. The adaptive nature of the V_{dc} can be noticed in the figure. As CPI voltage increases, consequently the V_{dc} increases and to feed the matching power to the grid, a decrease in inverter currents i_{abc} is observed at overvoltage.

V. Conclusion

An SPV system integrated with a distribution electrical grid network is modelled and simulated in MATLAB, and the performance of the system is analyzed. The system consists of two stages, one for boost converter and the other for VSC. The control technique proposed uses the d-q control approach and an adaptive DC-Link voltage according to the CPI variations, reducing losses and maintaining distortionless current injected into the grid. Furthermore, a grid-tied SPV system has been used, keeping in mind various problems related to the standalone systems. The proposed system proved to be much more flexible than the conventional constant DC-link voltage systems in times of voltage variations. The THD of the system were also found well below 5% (under IEEE-519 standards). The obtained simulation results have showcased the practicability of the proposed SPV system.

VI. Appendix

A. Parameters for simulation:

Interfacing Inductors- 1mH,

Ripple Filter- R=1 mΩ, and C=100μF,

PV Array- Peak Power-100KW, Open circuit voltage=363V, Short circuit current=368.48A

3 Phase grid- Voltage 415V, frequency=50Hz.

References

[1] J. T. Bialasiewicz, "Renewable Energy Systems With Photovoltaic Power Generators: Operation and Modeling," in *IEEE Transactions on Industrial Electronics*, vol. 55, no. 7, pp. 2752-2758, July 2008.

[2] R. Riestenberg and A. Singh, Efficiency of Solar Cells, Physics 141A April 23, 2013.

[3] MNRE (Ministry of New and Renewable Energy), Grid Connected Power/Solar. 2018. [Online]. Available: https://mnre.gov.in/solar

[4] P. Chiradeja, "Benefit of Distributed Generation: ALine Loss Reduction Analysis," *2005 IEEE/PES Transmission & Distribution Conference & Exposition: Asia and Pacific*, Dalian, China, 2005, pp. 1-5.

[5] A. Yadav and L. Srivastava, "Optimal placement of distributed generation: An overview and key issues," *2014 International Conference on Power Signals Control and Computations (EPSCICON)*, Thrissur, India, 2014, pp. 1-6.

[6] K. A. Joshi and N. M. Pindoriya, "Impact investigation of rooftop Solar PV system: A case study in India," *2012 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, Berlin, Germany, 2012, pp. 1-8

[7] H. Wang and D. Zhang, "The Stand-alone PV Generation System with Parallel Battery Charger," *2010 International Conference on Electrical and Control Engineering*, Wuhan, China, 2010, pp. 4450-4453

[8] B. Subudhi and R. Pradhan, "A Comparative Study on Maximum Power Point Tracking Techniques for Photovoltaic Power Systems," in *IEEE Transactions on Sustainable Energy*, vol. 4, no. 1, pp. 89-98, Jan. 2013.

[9] F. Liu, S. Duan, F. Liu, B. Liu and Y. Kang, "A Variable Step Size INC MPPT Method for PV Systems," in *IEEE Transactions on Industrial Electronics*, vol. 55, no. 7, pp. 2622-2628, July 2008.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:54:03 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3107857, IEEE Transactions on Power Delivery

Optimal Allocation of Capacitors in Radial Distribution Networks using Shannon's Entropy

Shubham Gupta, Vinod Kumar Yadav, *Senior Member, IEEE*, and Madhusudan Singh, *Senior Member, IEEE*

***Abstract*—Capacitor banks, when optimally placed in an electrical network improve power quality and reliability. For determining the capacitors' optimal location, the prevailing index methods mainly focus on single criterion, e.g., either of voltage stability, loss reduction, system loadability, and neglect some of the equally critical quantities while proving the solution. For obtaining a pragmatically feasible techno-commercial solution, it is advantageous to consider multiple criteria. This article proposes a novel approach for optimal capacitor allocation, which allows to choose as many objective functions (criteria) as desired. In the proposed approach, initially, the load flow study is executed to obtain the network's quantities for an uncompensated network. Subsequently, optimal locations of capacitor banks are identified through the proposed indexing method based on the application of Shannon's Entropy. The capacitor banks' optimal sizes are determined using the particle swarm optimization algorithm. The cost function is formulated as the minimization of sum of active power loss cost and the incurred annual capacitors cost subject to different operational constraints. The proposed approach is compared with that of the previously published methods by applying on radial distribution networks of different sizes (IEEE-12 bus, 34-bus and 108-bus) and the proposed approach is found to be superior.**

***Index Terms*—Load flow analysis, Radial Distribution network, Shannon's Entropy, optimal capacitor allocation, sensitivity index methods.**

I. NOMENCLATURE

The mathematical symbols and their abbreviations used in this article are as follows:

A. Sets

U_b: Set of buses in the network

Γ: Set of lateral branches in the network

U_{cap}: Set of optimal location for capacitor placement

B. Variables

$I_{line}(I_{bus})$: Branch (bus) current

$I_{line}^{real}\left(I_{line}^{imag}\right)$: Real (imaginary) component of line current

V_{bus}/V_{Li}: i^{th} bus voltage in l^{th} lateral

$V_s\ (V_r)$: Voltage at sending (receiving) end of the node

$P_{Li}(Q_{Li})$: Real (reactive) power at i^{th} sending end bus in l^{th} lateral

$P_e^{eff}\left(Q_r^{eff}\right)$: Effective real (reactive) power at receiving end node

P^{loss}: Total active power loss

$\bar{I}_{kl}$: Normalized index score of kth entity for l^{th} model

C. Parameters

λ_p: Active power loss cost per kW per annum

λ_{cap}: Capacitor installation cost per kVAr per annum

Q^T: Total demanded reactive power load

R: Line resistance in Ω

X: Line Reactance in Ω

Ω^{cap}: Capacitor size in kVAr

Q_t^{load}: Total demanded load

II. INTRODUCTION

GLOBALLY, the electrical networks are getting overstressed due to increasing demand for energy on the one hand and decreasing capacity of the existing networks on the other, because of various factors such as integration of distributed energy resources, shortage of reactive power, aging of extant infrastructure, and global warming, etc. This results in incidents of frequent blackouts, and huge monetary loss to society. Installation of shunt capacitor banks (SCBs) is one of the effective methods to destress the networks, apart from network expansion, installation of the distributed generators (DGs), voltage regulators (VRs), and FACTS devices [1], [2]. The root cause analysis reports on some of the recent blackout incidents [3], [4] indicate that deficit reactive power supply in the network was the leading cause behind blackouts.

Conventionally, the capacitors were installed near the primary substation; however, presently, the trend has changed. Installation of capacitors is now done based on the system requirements such as voltage profile improvement, loss reduction, increasing the system loadability by the installation of the capacitors at optimal locations. Numerous methods have been reported in the literature for optimal placement, sizing, type, or control scheme of capacitor banks [5]–[7].

Shubham Gupta is with Department of Electrical Engineering, Delhi Technological University, Delhi, India (shubham.gupta6691@gmail.com).

Vinod Kumar Yadav is with Department of Electrical Engineering, Delhi Technological University, Delhi, India (vinodkumar@dtu.ac.in).

Madhusudan Singh is with Department of Electrical Engineering, Delhi Technological University, Delhi, India (madhusudan@dce.ac.in).

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 09:14:23 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3107857, IEEE Transactions on Power Delivery

The limitation of computation capacity was the main reason behind earlier capacitor placement methods in the network, which were based on calculus and defined as an analytical approach. The first work carried out by [8] in 1956 proposed the famous '2/3 rule' for placement of the fixed capacitor in a uniformly distributed load in a distribution network. In [9] and [10], the methodology is extended by considering the switched capacitors or placement and sizing of capacitors on a uniform distribution network. Following these pioneering works, numerous analytical approaches have been proposed [5], [11], [12]. However, the analytical techniques required to make certain assumptions, such as uniformly distributed load, and these techniques also neglect the upfront investment of capacitors.

Later, the capacitor placement and sizing problem was formulated in [6] using dynamic programming, and a class of numerical programming approaches were evolved for capacitor allocation in distribution networks. The mathematical model optimized the objective function while satisfying certain constraints, e.g., voltage limit, capacitor size limit, and limit on the maximum number of capacitors. In [7], the capacitor placement problem was proposed as a mixed-integer programming problem. Integrating the heuristics in analytical or numerical programming approaches reduces the search space and convergence time to the optimal solutions [13]. Heuristic methods for capacitor placement and sizing were also reported by several other researchers [14]–[16].

Recently, artificial intelligence (AI) techniques have been implemented for capacitor placement and sizing problems, which dramatically reduced the search space and decreased overall computational time. In AI-based techniques, the genetic algorithm [17], particle swarm optimization (PSO) [18], fuzzy [19], flower pollination [13], and ant colony [20] are pretty powerful algorithms in solving complex problems and hence being used extensively. However, these AI methods require high memory space and computational time for solving complex systems. Recently, some authors have also dealt with the capacitor placement problem by combining the other network problems like voltage regulators, DGs placement, and load tap changer [21], [22].

In the modern power system, distributed energy resources (DERs) are gaining importance due to the major advantages like improving the voltage profile, power loss reduction, enhancing system security and reliability [23]. Moreover, there are other benefits of the DERs integration as below [24]:

- DERs share the load demands, reducing greenhouse gas and other air pollutants emissions by the central power plants.
- By local generation through DERs, line losses in transmission and distribution are reduced extensively and lesser energy production.
- In the case of fossil fuel-based power plants, warm water discharge, land usability, and other factors impact the environment directly or indirectly.

However, the optimal placement of DERs or DGs is very crucial from the system reliability and security aspects. The non-optimal placement of DGs can negatively impact the network, like an increase in power losses, weakening of the voltage profile. Many researchers proposed different approaches for the optimal placement of DGs [24], [25], in the past. However, the aspect of cost-effectiveness has been neglected in most of the solutions proposed in the past studies, which is important for distribution network operators (DNOs). It is worth to mention that DGs installation cost per MVA is much higher than capacitors, and hence it is wise to install the capacitors along with DGs to supply reactive power instead of installation of DGs [26].

The main objectives of optimal capacitor placement problem are to quantify the location, size, and type of the capacitors to be placed in a power system network to enhance the comprehensive techno-economic benefits. For determining the capacitors' optimal location, many index methods have been proposed in the past [20], [27], [28]. These methods identify potential nodes for capacitor placement based on identification of nodes having higher power loss, weak voltage magnitude, and other technical quantities. However, the prevailing index methods mainly focus on a single criterion, i.e., either of power loss reduction, voltage profile improvement, system loadability, while neglecting some of the equally critical quantities in the formulation. Also, while selecting an approach for optimal capacitors' placement arises questions on the subjectivity of choice of a particular criterion.

This article aims at investigating a novel holistic approach for finding the capacitors' optimal location and size in a radial distribution network (RDN), with the objective to reduce the active power losses and improve the voltage profile. The proposed approach is based on the application of Shannon's Entropy (SE), defined as the Shannon Entropy-based index method (SE-IM). The concept of Shannon's entropy effectively handles the drawbacks of prevailing methods, mentioned in paragraph above by syndicating all the desired criteria based on their degree of importance [29]. The SE-IM determines the optimal location for siting switched and (or) fixed capacitors. In the second stage, the classical Particle Swarm Optimization (PSO) algorithm is applied to size capacitors at identified optimal locations. The efficacy of the proposed work is evaluated by implementing the proposed approach on an IEEE-12 bus, 34 bus and a practical 108 bus RDN of the Indian utility system, and the analysis of results show promising outcomes. The proposed methodology has the following main advantages and novelty:

- A direct approach for load flow analysis evaluates the network's quantities in a given RDN, which required less computational time to converge even for large networks.
- Novel Shannon's Entropy-based index method (SE-IM) allows considering many objective functions (criterion) as desired. SE-IM approach combines these criteria based on their degree of importance defined by the network operator.
- The SE-IM approach is being applied to eliminates the subjectivity on the choice of a criterion to quantify the optimal capacitors' location.
- The classical PSO algorithm is utilized to solve the sizing problem of capacitors, which is computationally efficient and easy to implement compared to the other heuristic algorithms.

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 09:14:23 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3107857, IEEE Transactions on Power Delivery

The rest of the paper is organized as follows. Section III describes the problem formulation and the approach for load flow analysis. Section IV delineates the Shannon's Entropy approach applied for raking the potential solutions of optimal allocation and presents the proposed SE-IM approach for indexing the potential buses for optimal siting the capacitors. Section V presents the simulated results, followed by concluding remarks in Section VI.

III. Optimal Capacitor Allocation Problem Formulation

A. Direct Approach for Load Flow Analysis

In the present analysis, at first a load flow analysis is carried out to obtain base case (i.e., without compensation) active power flow and voltage profiles through the direct approach of load flow analysis. This approach is more suitable for radial networks than the classical load flow analysis techniques and also requires less computational time to converge [30]. It evaluates two matrices using topological network attributes, which are bus injection to branch current (BIBC) and branch current to bus voltage (BCBV) matrices. These matrices relate bus current to the bus voltage through branch current in the network. For a 3-bus RDN (see Fig. 1), the relationship between the bus current, branch current, and bus voltage may be expressed by (1) and (2):

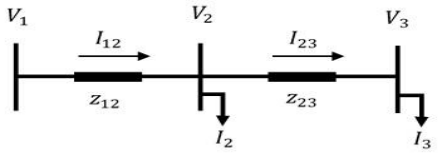


Fig. 1. 3-bus network depicting bus and line parameters.

$$\begin{bmatrix} I_{12} \\ I_{23} \end{bmatrix}_{2\times1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}_{2\times2} \begin{bmatrix} I_2 \\ I_3 \end{bmatrix}_{2\times1} \tag{1}$$

$$\begin{bmatrix} V_1 \\ V_1 \end{bmatrix}_{2\times1} - \begin{bmatrix} V_2 \\ V_3 \end{bmatrix}_{2\times1} = \begin{bmatrix} z_{12} & 0 \\ z_{12} & z_{23} \end{bmatrix}_{2\times2} \begin{bmatrix} I_{12} \\ I_{23} \end{bmatrix}_{2\times1} \tag{2}$$

The bus voltage (V_1) is a slack bus in (2) and its magnitude is equal to 1. The voltages at bus 2 and 3 in (2) can be quantified using the branch currents I_{12} and I_{23}, which are determined by (1) using the bus currents I_2 and I_3. In general, for a main feeder radial distribution network having $\{1,2,\ldots,n\} \in U_b \setminus N^o$ buses; the formulation of the BIBC and BCBV matrices may be represented by (3) and (4), respectively.

$$\left(I^k_{line}\right)_{(n-1)\times1} = [BIBC]_{(n-1)\times(n-1)} \left(I^k_{bus}\right)_{(n-1)\times1} \tag{3}$$

In (3), BIBC is a $((n-1)\times(n-1))$ matrix that dictates the relationship between bus and branch currents at k[th] iteration. Similarly, the BCBV is a $((n-1)\times(n-1))$ matrix that portrays the relationship between the branch currents and bus voltages in the RDN.

$$\left(V_1 - V^k_{bus}\right)_{(n-1)\times1} = [BCBV]_{(n-1)\times(n-1)} \left(I^k_{line}\right)_{(n-1)\times1} \tag{4}$$

From (3) and (4), the distribution load flow matrix can be formed as

$$\left(V_1 - V^k_{bus}\right)_{(n-1)\times1} = [BIBC]_{(n-1)\times(n-1)} [BCBV]_{(n-1)\times(n-1)} \left(I^k_{bus}\right)_{(n-1)\times1} \tag{5}$$

where

$$[BIBC]_{(n-1)\times(n-1)} [BCBV]_{(n-1)\times(n-1)} = [DLF]_{(n-1)\times(n-1)} \tag{6}$$

Several load flow studies have utilized this approach to carry out the load flow analysis, e.g., [31], [32], and found to be superior as the method saves computational time by avoiding LU factorization and backward-forward swapping of the Y admittance matrix or Jacobian matrix. Instead, it exploits the topological characteristics of RDN to solve load flow problems directly [30].

B. Minimization of Cost of Annual Energy Loss

This proposed approach focuses on the minimization of cost of annual energy loss (AEC) in $/year, which is constituted by the annual active power loss cost ($/kW) and the capital investment cost ($/kVAr) of siting the capacitors at potential locations.

Mathematically, the AEC (in $) function $S(\mathrm{P}^{loss}, \Omega^{cap})$ can be presented as:

$$S(\mathrm{P}^{loss}, \Omega^{cap}) = \left[\lambda_p \sum_{L\in\Gamma} \sum_{i\in U_b\setminus\{n\}} R_{Li+1} \frac{P^2_{Li} + Q^2_{Li}}{V^2_{Li}} \right] + \left[\lambda_{cap} \sum_{j\in U_{cap}} \Omega^{cap}_j \right] \tag{7}$$

where λ_p and λ_{cap} are the cost parameters for active power loss cost ($/kW/year) and capacitor installation cost ($/kVAr/year), respectively. The expected life considered for a capacitor bank is ten years. The capacitor allocation problem can be formulated as a minimum of AEC function subjected to line flow and variable bound constraints, given by (8) -(12):

$$\min_{[P,Q,V,\Omega]} . \; S(\mathrm{P}^{loss}, \Omega^{cap}) \tag{8}$$

Subject to:

$$\left(V_1 - V_{bus}\right) + [DLF](I_{bus}) = 0 \tag{9}$$

$$U_{cap} - U^{\max}_{cap} \le 0 \tag{10}$$

$$\Omega^{cap}_t - \alpha.Q^{load}_t \le 0 \tag{11}$$

$$\Pi(V_{bus}, \Omega^{cap}) \le 0 \tag{12}$$

where (8) represents the annual energy cost ($/year) as an objective function, (9) dictate the load flow analysis explained earlier, (10) represents the limit on capacitors number $U^{\max}_{cap}$ that can be installed, (11) represents the bound on total maximum capacitor size $\alpha.Q^{load}_t$ that can be installed and (12) represents the bounds on voltage and individual capacitor size. The value of α in (11) is selected so that the power factor will remain lagging or unity. However, due to intermittent capacitor sizes in a fixed type capacitor scheme, it may be possible that the power factor becomes marginally leading.

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 09:14:23 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3107857, IEEE Transactions on Power Delivery

The minimum and maximum capacitor sizes (range) based on the two type schemes (switched or fixed) considered in this article are:

$$\Omega^{cap} = \begin{cases} (0-1200)\ kVAr; & \text{switched type capacitor} \\ \begin{bmatrix} 150, 300, 450, \\ 600, 750, 900, \\ 1050, 1200 \end{bmatrix} kVAr; & \text{fixed type capacitor} \end{cases}$$

To determine the optimal size of the capacitors in the considered RDN, classic PSO is applied in the present work [33], as this algorithm requires lesser computational resources and time and also easy to implement compared to other AI-based algorithms. In order to obtain the optimal sitting of the capacitors, the novel approach based on the SE is evaluated in this article, which is discussed in the following Section.

IV. Solution Methodology

The prevailing index methods for optimal allocation of capacitor are mainly focused on a single criterion, which is either of reduction of losses in the line or improvement of voltage at buses or increasing the loadability of the network. This article presents a novel methodology that enables the distribution network operator to include as many criteria as desired at a time for optimal capacitor allocation. In this approach, the optimal capacitor allocation problem is solved in two stages. First, a load flow study is conducted for given busloads and line data, termed as the base case (un-compensated), to get the active power flow in each line and voltage profiles at the buses. In the second stage, the optimal site to place the capacitor is determined using the SE; and optimal size is computed through classic PSO and given in Appendix. The following sub-section briefly describes the concept of Shannon's Entropy

A. Modeling of Shannon's Entropy (SE)

The concept of Shannon's Entropy was proposed to quantify the degree of randomness in a variable in information theory by C.E. Shannon in 1948 [34]. This paper utilizes the principle to acquire the degree of importance of an indexing model to obtain a cumulative index score. The SE formula e_l can be represented as:

$$e_l = -e_0 \sum_{k=1}^{s} \bar{\mathrm{I}}_{kl} . \ln \bar{\mathrm{I}}_{kl} \tag{13}$$

where e_0 denotes the entropy constant and is equal to $(\ln s)^{-1}$, $\bar{\mathrm{I}}_{kl}$ represents the normalized index scores of kth entity (i.e., bus in this case) for all considered indexing models (l) in the formulation. Furthermore, the degree of diversification can be obtained as: $\delta_l = 1 - e_l$; by normalizing $\mu_l = \delta_l \big/ \sum_{l=1}^{r} \delta_l$, where μ_l is the degree of importance of an index model. The cumulative index score of any kth entity is evaluated as:

$$\eta_k = \sum_{k=1}^{s} \mu_l I_{kl} \tag{14}$$

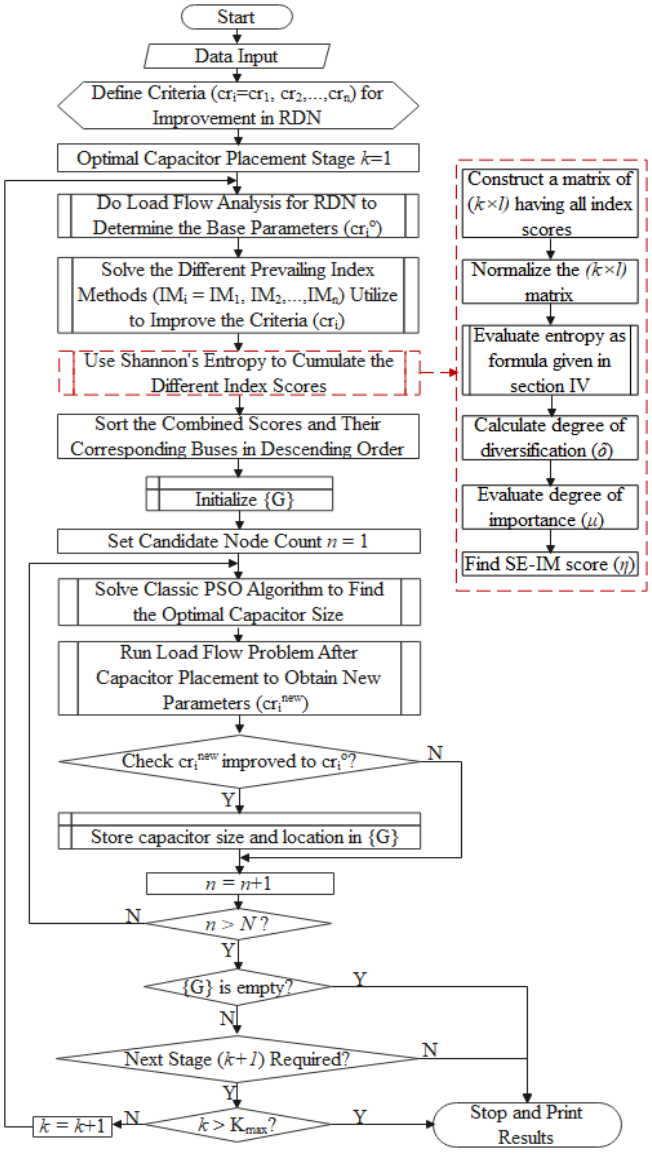


Fig. 2. The workflow of the proposed SE-IM approach.

B. Proposed Shannon's Entropy-based Index method (SE-IM)

After the formulation of load flow study for an un-compensated network, SE-IM prioritizes the buses based on their index scores, which is computed by combining three existing indices, which are loss sensitivity index (LSI) [20], power loss index (PLI) [27], and voltage sensitivity index (VSI) [28]; a parametric study is summarized in Table I. This approach allows DNO to incorporate many criteria at a time to emulate the real-world scenario. The central aid of using SE is to define the degree of importance for each index method (or criteria). The steps involved in solving the optimal capacitor problem using novel SE-IM approach are described in the flow chart of the algorithm is shown in Fig. 2.

1. Data mining: load the bus data and line data of the network under study.
2. Identify or define the criteria required for improvement in RDN.
3. Optimal capacitor placement stage k = 1.

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 09:14:23 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3107857, IEEE Transactions on Power Delivery

TABLE I
EXISTING INDEX METHODS FOR OPTIMAL CAPACITOR PLACEMENT

Index Methods	Formulae	Feature/ Comment
LSI_1	$LSI_1 = \frac{2 \times Q_r^{eff} \times R_{line}}{\lvert V_r \rvert^2}$	- Obtained through the first derivate of P_{loss} with respect to the Q_r.
LSI_2	$LSI_2 = -2 \times R_{line}\left[\frac{\left(P_r^{eff}\right)^2 + \left(Q_r^{eff}\right)^2}{\lvert V_r \rvert^3}\right]$	- Obtained through the first derivate of P_{loss} with respect to the V_r.
Vector Index (PLI)	$PLI = \frac{1}{\lvert V_r \rvert^2} + \frac{I_{line}^{imag}}{I_{line}^{real}} + \frac{Q_r^{eff}}{Q^T}$	- Both capacitor location and size can be evaluated analytically.
VSI	$VSI = \lvert V_s \rvert^4 - 4 \times \left(P_r X_{line} - Q_r R_{line}\right)^2 \cdots$ $\cdots - 4 \times \left(P_r R_{line} + Q_r X_{line}\right) \times \lvert V_s \rvert^2$	- Identify node which most sensitive to voltage collapse.

4. Do the load flow analysis as given in Section III for the base case (i.e., uncompensated RDN).
5. Solve different existing index methods which improve the criteria mentioned in STEP 2.
6. Combine the index scores using the SE described in Section IV.
7. Sort the SE-IM score and their respective bus in descending order. The buses with high scores get priority for siting of capacitors.
8. Initialize an empty set {G} and set the candidate node count n = 1.
9. Apply the classic PSO algorithm to evaluate the optimal capacitor size Ω^{cap}, and repeat load flow analysis after capacitor placement.
10. Check whether the criteria obtained in STEP 9 are improved than the obtained in STEP 4. If yes, store the candidate node and capacitor size in {G}. Increment the count n by 1.
11. If n > max (N): go to STEP 12; else, go to STEP 9.
12. If {G} is an empty set: go to STEP 14; else, go to STEP 13. Empty {G} set signifies that no further significant improvement can be gained with further addition of capacitor, so algorithm stops.
13. If the next capacitor placement stage is to be carried, then increment k by one and check k < max (K); if yes, go to STEP 4.
14. Stop and print the obtained results.

V. RESULTS AND DISCUSSION

The simulations study are carried out on radial distribution networks for IEEE 12 bus, 34 bus and 108-bus practical RDN of Indian utility to demonstrate the efficacy of the proposed approach. The comparison of the proposed approaches has been made with four prevailing index methods, which are loss sensitivity index methods 1 and 2 (LSI_1&2), vector-index method (PLI), and voltage sensitivity index method (VSI). The cost parameters (λ_p and λ_{cap}) are considered as 168 \$/kW/annum and 5 \$/kVAr, respectively. The base quantities are considered for load flow studies are: $S_{base} = 100 MVA$ and $kV_{base} = 11 kV$. Table II reports the load flow solutions for base (before capacitor placement) RDNs at one p.u. load level. In table II, the computational time (including the time elapsed in solving the load flow problem, data/results calculations, and display of results) to converge the load flow solutions is stated. It is noteworthy that, as the network size increases, the considered load flow method becomes more time-efficient.

For the compensation of network, two cases based on the capacitor type have been considered; switched and fixed capacitor type schemes; the ranges or available sizes for capacitors are mentioned in Section II. After determining the optimal siting, the optimal capacitor size is evaluated using the classic PSO algorithm. The result of the above approach is delineated in subsequent paragraphs.

TABLE II
LOAD FLOW ANALYSIS FOR UNCOMPENSATED NETWORKS (BEFORE CAPACITOR PLACEMENT)

Test system	P-Loss (MW)	V_{max}/V_{min} (p.u.)	$\lvert I_{1\text{-}2} \rvert$ (p.u.)	Weakest voltage bus	Computational time (sec.)	AEC (\$)
12-Bus	0.0207	1 / 0.9434	0.0062	12	0.022	3479.28
34-Bus	0.2217	1 / 0.9417	0.0568	27	0.025	37248.94
108-Bus	0.6450	1 / 0.8944	0.0527	105	0.041	108363.14

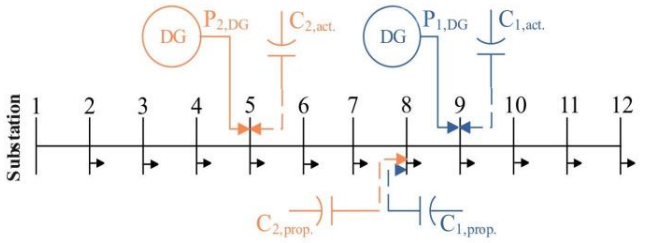


Fig. 3. Single line diagram of IEEE-12 bus RDN including a distributed generator in case-1 (blue) and case-2 (orange).

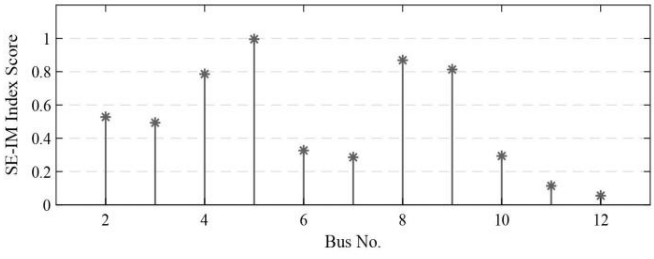


Fig. 4. SE-IM index score for 12 bus RDN at 1 p.u. load level.

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 09:14:23 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3107857, IEEE Transactions on Power Delivery

TABLE III
COMPARISON AMONG DIFFERENT EXISTING INDEX METHODS AND PROPOSED APPROACH FOR IEEE-12 BUS RDN

Method	Capacitor Position / Size (MVAr)	P-Loss (MW)	V_{max}/V_{min} (p.u.)	AEC ($)	Net Saving (%)
Method 1: Loss Sensitivity Index (1)					
Case-1	8 / 0.300	0.0129	1 / 0.9574	2169.52	33.33
Case-2	9 / 0.1284	0.0138	1 / 0.9515	2317.68	31.54
Method 2: Loss Sensitivity Index (2)					
Case-1	5 / 0.300	0.0141	1 / 0.9518	2365.37	27.70
Case-2	5 / 0.2333	0.0145	1 / 0.95	2437.85	26.58
Method 3: Vector Index (PLI)					
Case-1	4 / 0.450	0.0163	1 / 0.9506	2735.33	14.92
Case-2	4 / 0.4107	0.0161	1 / 0.95	2698.72	16.53
Method 4: Voltage Sensitivity Index (VSI)					
Case-1	12 / 0.150	0.0136	1 / 0.9545	2283.71	32.21
Case-2	12 / 0.1110	0.0144	1 / 0.9518	2417.59	28.92
Method 5: Shannon's Entropy based Index Method (Proposed)					
Case-1	8 / 0.300	0.0129	1 / 0.9574	2169.52	33.33
Case-2	9 / 0.1284	0.0138	1 / 0.9515	2317.68	31.54

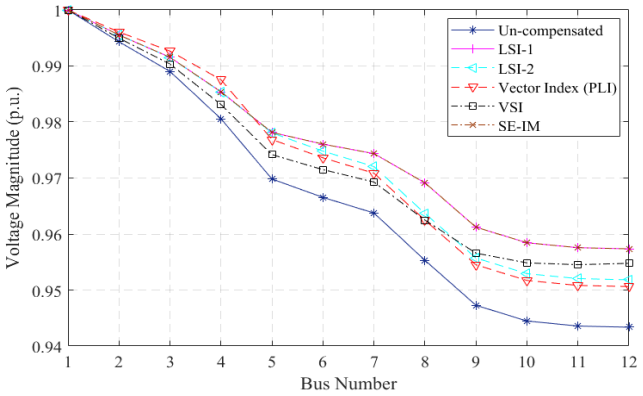


Fig. 5. Voltage profiles for IEEE-12 bus RDN in case-1 (fixed capacitors).

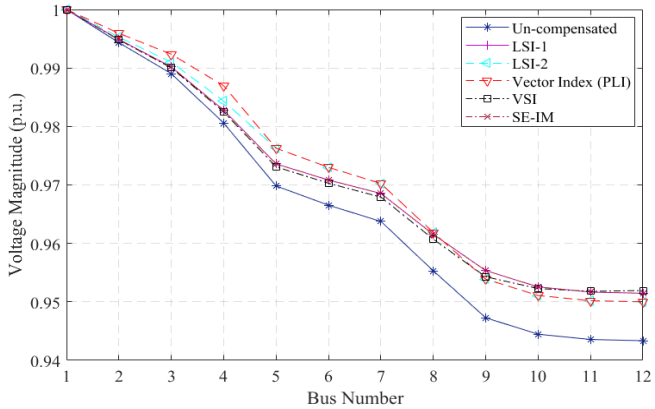


Fig. 6. Voltage profiles for IEEE-12 bus RDN in case-2 (switched capacitors).

A. Results of IEEE-12 Bus Radial Distribution Network

The IEEE-12 bus RDN is structurally simple and small network consisting of 11 lines connecting 12 buses radially as shown in Fig.3. The total active and reactive powers supplied from the substation are 0.456 MW and 0.413 MVAr, respectively. The total active power loss that occurred in the RDN is 0.0207 MW as shown in Table II. The SE-IM index score is quantified for all the nodes at one p.u. load level and presented in Fig. 4. A comparative analysis between the prevailing index approaches and the proposed approach for the optimal allocation and sizing of the capacitor is performed and results are given in Table III. For simulations of these approaches, it was considered that the upper and lower voltage limits are bounded to be 0.95 and 1.05 p.u., respectively. The maximum number of locations (U_{cap}^{max}) where the capacitor can be placed is restricted to one in each case. In table III, case-1 and case-2 represent the fixed capacitor type and switched capacitor type scheme, respectively. Followings are the main findings of the simulation study.

1) After the capacitor placement, there are significant improvements in techno-economical quantities that can be seen in comparison to the base case. The major technical quantities considered in the present analysis are the power loss reduction, voltage profile improvements, and reduction in the network's current (or power) flows. In addition, the analysis of economic quantity is the annual cost incurred due to the active power loss.
2) The loss sensitivity index (Method 1) and the proposed approach prioritized the same buses for optimal capacitor placement, *i.e.*, 8th bus for fixed capacitor and 9th bus for switched capacitor type schemes.
3) The SE-IM approach mainly improves all the techno-economic quantities as compared to other method, except the node voltages in Case 2, which is significantly improved by the VSI (Method 4). The voltage profiles are delineated in Figs. 5-6.
4) The results show that node 8 is best suited for a fixed capacitor placement scheme while node 9 is appropriate for a switched capacitor placement scheme. Finally, the PLI (Method 3) shows node 4 for optimal capacitor placement and generating the least preferable results in all manner.

Discussion on the capacitor size is omitted as this work mainly focuses on determining the optimal location for capacitor placement. However, the capacitor size is determined to enable the techno-economical comparison between the prevailing methods and the proposed method.

1) Optimal Capacitor Placement in the Presence of Distributed Energy Resources

An analysis is carried out to understand the effectiveness of the proposed approach in the presence of DERs. The computed results only consider the fixed type capacitor placement for the sake of simplicity and presented in Table IV. It is assumed that the distributed generator (DG) is of type-1 (supply only real power P_{DG}), and two cases based on the location of DG are considered as follow (see Fig.3):

Case-1. When DG of 0.235 MW is placed at the 9th bus.

Case-2. When DG of 0.280 MW is placed at the 5th bus.

It is observed that the total real power loss reduction is more using the SE-IM approach for both cases (i.e., 83.09% and 73.43%, respectively) in comparison to the actual scenario (DG and capacitor are placed at the same node). The voltage profile improvement and real power flow curves are shown in Fig. 7 and Fig. 8, respectively.

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 09:14:23 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3107857, IEEE Transactions on Power Delivery

TABLE IV
COMPARISON BETWEEN TWO METHODS FOR ADDING SINGLE CAPACITOR IN THE PRESENCE OF DERS

Method	DG Location / Size in MW	Capacitor Position / Size in MVAr	P-Loss (MW)	V_{max}/V_{min} (p.u.)	Net P-loss Reduction (%)
Actual*					
Case-1	9 / 0.235 [35]	9 / 0.150	0.0038	1 / 0.9889	81.64
Case-2	5 / 0.280 [35]	5 / 0.300	0.0066	1 / 0.9703	68.12
Proposed Approach					
Case-1	9 / 0.235	8 / 0.300	0.0035	1.0007 / 0.9934	83.09
Case-2	5 / 0.280	8 / 0.300	0.0055	1 / 0.9757	73.43

* It is assumed to place DG and capacitor at the same location.

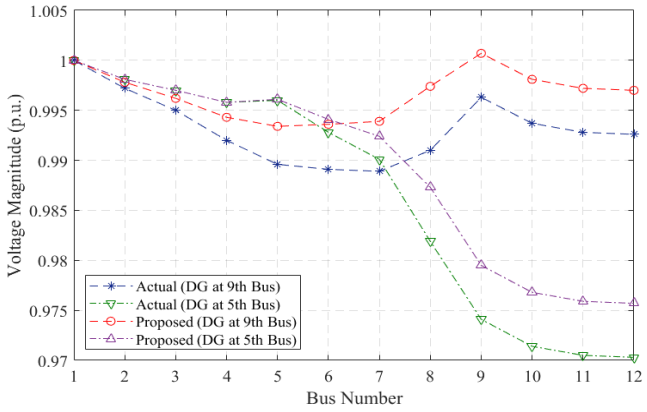


Fig. 7. Voltage profile for IEEE-12 bus RDN after capacitor placement in the presence of single DG.

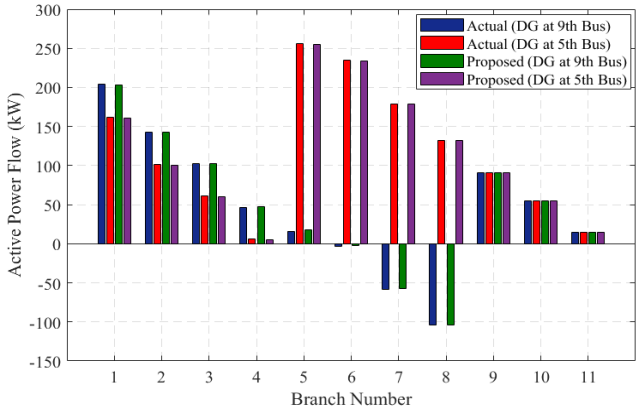


Fig. 8. Active power flow in IEEE-12 bus RDN after capacitor placement in the presence of single DG.

B. *Result of 34 Bus Radial Distribution Network*

The proposed approach synergizes four prevailing methodologies through Shannon's entropy. The comparison of the result of the proposed approach vis-à-vis that of the prevailing index methods is presented in Table V. The load flow results for the base network are given in Table II, where total active loss was 221.72 kW and total annual loss cost incurred was USD 37248.94. For simulations study using these approaches, it was considered that the maximum number of locations (U_{cap}^{max}) where the capacitors can be placed is limited to four in each case, and the minimum and maximum voltages limits on the buses are considered to be 0.95 p.u. and 1.05 p.u., respectively. In (11), the value of α is assumed to be 1.1, *i.e.*, the total kVAr capacity that can be injected into the network is limited to 1.1 times the total demand (i.e., 2873.5 kVAr). In Table V, case-1 represents the fixed capacitor placement, and case-2 represents the switched capacitor placement in RDN. Some of the important findings of this analysis are discussed below.

1) It is observed that there is a reduction in active power loss, and annual saving in $/year and optimum for the proposed approach compared to the existing index methods. Though in case-1, the active power loss reduced more by VSI method and voltage profile is better for vector-index (PLI) method; however, the combined active power loss and voltages profile is significantly improved for the SE-IM approach. The minimum and maximum voltage magnitudes are tabulated in Table V as well.
2) The annual energy loss cost (AEC) is minimum for the proposed approach in case-2 (switched capacitor scheme), i.e., USD 28252.64.
3) Figs. 9 and 10 show the voltage profiles before and after compensation in case-1 and case-2, respectively. Fig. 11 depicts the potential locations for capacitors allocation in 34 bus RDN, red color represents the switched capacitors, and green color represents the fixed capacitors.
4) The optimal locations and capacitors' sizes computed for the proposed approach is presented in Table VI.

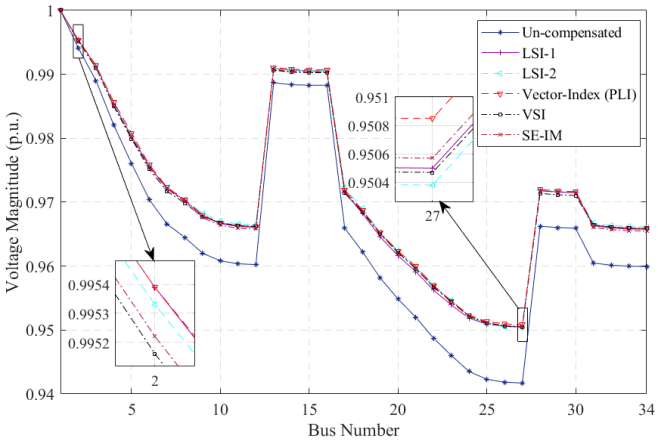


Fig. 9. Voltage profiles for 34 bus RDN in case-1 (fixed capacitors).

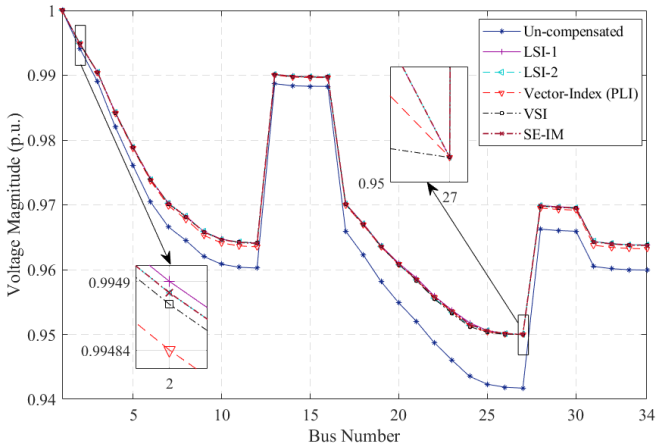


Fig. 10. Voltage profiles for 34 bus RDN in case-2 (switched capacitors).

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 09:14:23 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3107857, IEEE Transactions on Power Delivery

TABLE V
COMPARISON AMONG DIFFERENT EXISTING INDEX METHODS AND PROPOSED APPROACH FOR 34-BUS RDN

Method	Total Capacitor Size (MVAr)	P-Loss (MW)	V_{max}/V_{min} (p.u.)	AEC ($)	Net Saving (%)
Method 1: Loss Sensitivity Index (1)					
Case-1	3.000	0.1624	1 / 0.9505	27285.32	22.72
Case-2	1.7773	0.1688	1 / 0.95	28350.90	21.50
Method 2: Loss Sensitivity Index (2)					
Case-1	2.850	0.1613	1 / 0.9504	27095.32	23.43
Case-2	1.7620	0.1688	1 / 0.95	28356.70	21.51
Method 3: Vector Index (PLI)					
Case-1	3.000	0.1610	1 / 0.9509	27040.14	23.38
Case-2	1.6261	0.1716	1 / 0.95	28820.98	20.44
Method 4: Voltage Sensitivity Index (VSI)					
Case-1	2.400	0.160	1 / 0.9505	26874.67	24.63
Case-2	1.7221	0.1684	1 / 0.95	28295.59	21.72
Method 5: Shannon's Entropy based Index Method (Proposed)					
Case-1	2.550	0.1609	1 / 0.9506	27037.32	23.99
Case-2	1.7472	0.1682	1 / 0.95	28252.64	21.81

TABLE VI
OPTIMUM LOCATION AND SIZE IN MVAR OBTAINED USING PROPOSED APPROACH

Case	Optimal Location and Size in MVAr
Case-1	19 (0.600); 7 (1.050); 23 (0.450); 25 (0.450)
Case-2	20 (0.2176); 23 (0.4592); 9 (0.2261); 25 (0.8444)

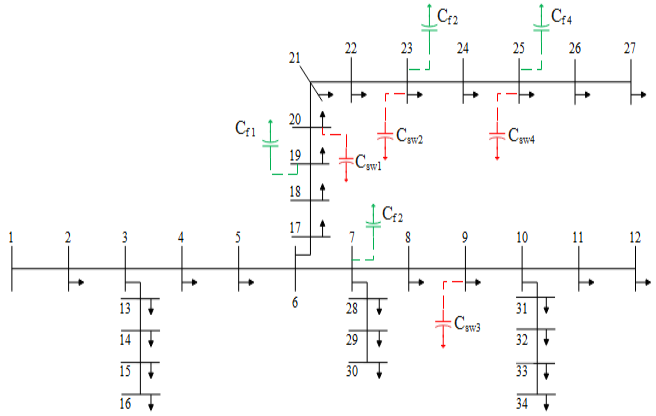


Fig. 11. 34 bus RDN depicting potential locations for switched capacitors (red) or fixed capacitors (green) using SE-IM approach.

C. Result of 108 bus RDN of an Indian utility

The comparative results for 108-bus practical RDN of an Indian utility are reported in Table VII. The maximum number of locations where the capacitors can be placed is limited to 15 in each case. The minimum and maximum voltages on the buses are considered as 0.90 p.u. and 1.10 p.u., respectively. In (11), the value of α is assumed to 1, which implies that the maximum total kVAr capacity that can be injected into the network is considered to be equal to the total demanded load (i.e., 9.099 MVAr). Table VII shows that active power loss reduced significantly for the proposed approach compared to that of the prevailing methods; as a result, the AEC noticeably reduced. In Table VII, case 1 and case 2 represent the fixed and switched capacitor placement type schemes in RDN.

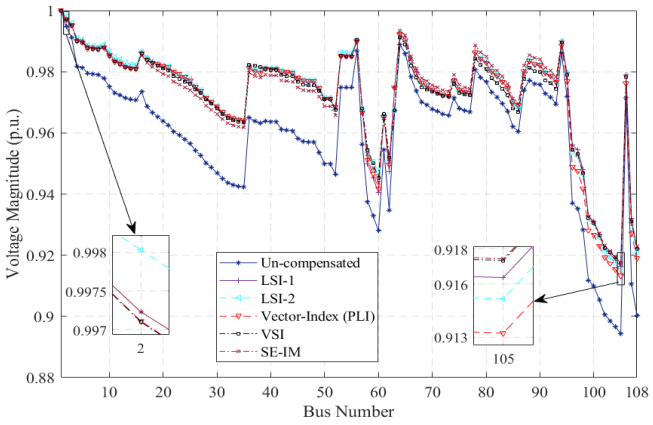


Fig. 12. Voltage profiles for practical 108 bus Indian RDN in case-1.

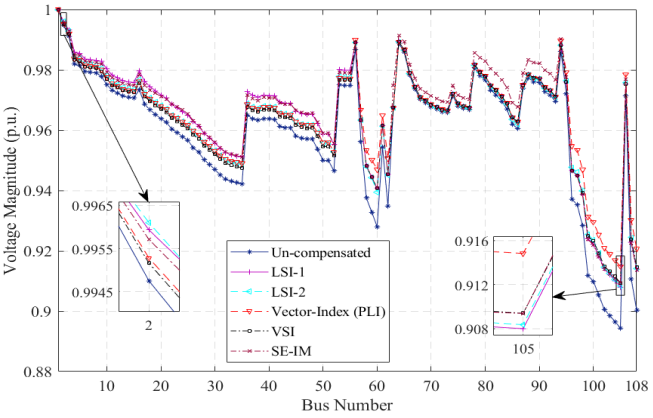


Fig. 13. Voltage profiles for practical 108 bus Indian RDN in case-2.

TABLE VII
COMPARISON AMONG DIFFERENT EXISTING INDEX METHODS AND PROPOSED APPROACH FOR 108-BUS RDN

Method	Total Capacitor Size (MVAr)	P-Loss (MW)	V_{max}/V_{min} (p.u.)	AEC ($)	Net Saving (%)
Method 1: Loss Sensitivity Index (1)					
Case-1	8.850	0.4363	1 / 0.9164	73302.56	28.27
Case-2	3.1958	0.4954	1 / 0.9080	83218.08	21.73
Method 2: Loss Sensitivity Index (2)					
Case-1	8.700	0.4485	1 / 0.9152	75353.64	26.45
Case-2	3.4878	0.5108	1 / 0.9084	85805.31	19.21
Method 3: Vector Index (PLI)					
Case-1	7.800	0.4356	1 / 0.9132	73186.85	28.86
Case-2	3.4793	0.4935	1 / 0.9148	82903.68	21.89
Method 4: Voltage Sensitivity Index (VSI)					
Case-1	6.750	0.4630	1 / 0.9173	77777.22	25.11
Case-2	2.2140	0.5082	1 / 0.9094	85383.18	20.18
Method 5: Shannon's Entropy based Index Method (Proposed)					
Case-1	9.000	0.4227	1 / 0.9175	71012.79	30.32
Case-2	6.0445	0.4699	1 / 0.9094	78942.33	24.36

The SE-IM approach gives flexibility to the network operators for selecting the various criterion based on requirements. The inclusion of as many criteria into the formulation will make the optimal capacitor allocation problem more realistic. Fig. 12 and Fig. 13 present the voltage profiles calculated through the prevailing and the proposed methods.

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 09:14:23 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3107857, IEEE Transactions on Power Delivery

TABLE VIII
RESULTS OF OPTIMAL CAPACITOR PLACEMENT PROBLEM FOR PRACTICAL 108 BUS INDIAN RDN IN THE PRESENCE OF LOAD UNCERTAINTY

Load Level	Before compensation		After compensation			Net AEC saving (%)
	P-Loss (MW)	V_{max}/V_{min} (p.u.)	Total Capacitor Size (MVAr)	P-Loss (MW)	V_{max}/V_{min} (p.u.)	
Case-1: Fixed capacitor type scheme						
Light	0.1515	1 / 0.9499	4.500	0.1037	1 / 0.9571	22.71
Nominal	0.6450	1 / 0.8944	9.000	0.4227	1 / 0.9175	30.32
Peak	1.8028	1 / 0.8177	16.650	1.6073	1 / 0.9001	8.10

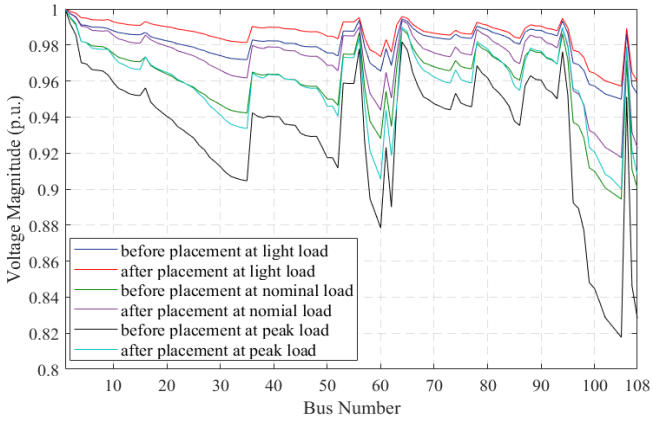


Fig. 14. Voltage profiles for practical 108 bus Indian RDN in the presence of load uncertainty.

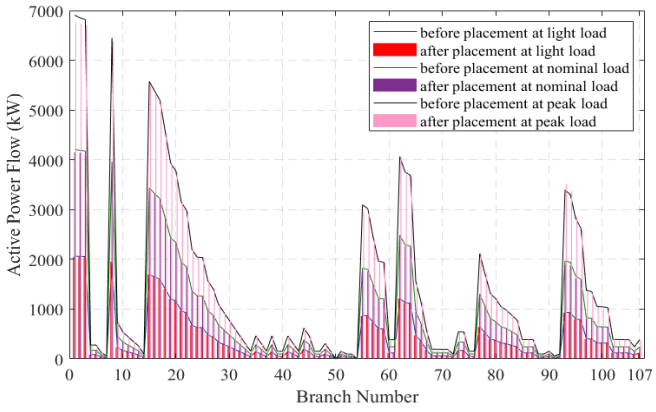


Fig. 15. Active power flows in practical 108 bus Indian RDN in the presence of load uncertainty.

1) Optimal Capacitor Placement in the Presence of Load Uncertainty

To examine the effectiveness of the proposed SE-IM approach in the presence of load uncertainty, three load levels are considered, *i.e.*, light, nominal, and peak, and their respective load multiplying factors are 0.5, 1.0, and 1.6. The results are obtained for fixed capacitor type scheme, reported in Table VIII. Figs. 14 and 15 show the voltage profiles and active power flow profiles before and after the compensation during light, nominal and peak load conditions, respectively. Followings are the main findings from the above study.

1) At light load condition, the net real power loss is reduced by 31.55%, while the AEC saving is 22.71% after the placement of capacitors of total capacity 4.50 MVAr.
2) At nominal load condition, the power loss reduction and AEC saving are optimum than light and peak load condition.
3) AEC savings are less in peak load conditions. It is because of the limitation on the total number of places where capacitors could be placed, same as the nominal or light load condition.

VI. CONCLUSIONS

The main focus of this paper is to present a novel approach that allows the network operators to select multiple criteria in the formulation at a time for optimal capacitor siting. The efficacy of the proposed approach is evaluated by implementing it on an IEEE-12 bus, 34 standard bus and a 108-bus practical RDN of an Indian utility, and the result is compared with that of the prevailing four methods and the proposed method is found to be superior. The significant contributions of the article are summarized below.

1) In this article, two cases of analysis have been considered based on the capacitor types, which are switched and fixed types.
2) The proposed approach comprehensively improved the techno-economic quantities, which fulfilled the objectives of the optimum siting problem effectively.
3) An impact on optimal capacitor placement problem is analyzed in the presence of DERs and load uncertainty in the RDN. It is noted that; the proposed approach improved the network's quantities while satisfying all the considered security and reliability constraints.
4) The proposed SE-IM approach offers flexibility in selecting multiple criteria, and also enables combining the index scores obtained from the methods considered based on their degree of importance and thereby improves the network quantities.

APPENDIX

Particle Swarm Optimization (PSO) is a stochastic technique to find the optimum solution based on the socio-psychological behavior of birds or fish. The algorithm starts with a number of random solutions of a particle i as $x_i = \{x_{i1}, x_{i2}, x_{i3},...,x_{ij}\}$ and updates after each iteration k to get the best solution achieved by i^{th} particle so far as $\rho_i^{best} = \{\rho_{i1}^{best}, \rho_{i2}^{best}, \rho_{i3}^{best},..., \rho_{ij}^{best}\}$. Concurrently, a best solution so far attained by any particle is also recorded, defined as the global best solution (ψ_{ij}^{best}). After each iteration, the velocity $v_i = \{v_{i1}, v_{i2}, v_{i3},...,v_{ij}\}$ and solution x_{ij} of the ith particle are updated by using following equations:

$$v_{ij}^{k+1} = \omega \times v_{ij}^{k} + \left[c_1 \times \alpha_1 \times \left(\rho_{ij}^{best} - x_{ij}\right)\right] + \left[c_2 \times \alpha_2 \times \left(\psi_{ij}^{best} - x_{ij}\right)\right] \quad \text{(A-1)}$$

$$x_{ij}^{k+1} = x_{ij}^{k} + v_{ij}^{k+1} \quad \text{(A-2)}$$

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 09:14:23 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3107857, IEEE Transactions on Power Delivery

where k is the iteration count, c_1 and c_2 are the constants and equal to 2, α_1 and α_2 represents the random variables and ω represents the inertia weight which modifies after every iteration as:

$$\omega = \omega^{\max} - \left[\left(\omega^{\max} - \omega^{\min} \right) \times \frac{k}{\mathrm{K}^{\max}} \right] \quad \text{(A-3)}$$

where $\omega^{\max}$ and $\omega^{\min}$ are the maximum and minimum inertia weights and equal to 0.9 and 0.4, respectively. In this article, the total 100 number of iterations ($\mathrm{K}^{\max}$) are carried out for each case to obtain the optimal solution.

REFERENCES

[1] N. K. Meena, A. Swarnkar, N. Gupta, and K. R. Niazi, "Optimal integration of DERs in coordination with existing VRs in distribution networks," *IET Gen. Tran. Dist.*, vol. 12, no. 11, pp. 2520–2529, 2018.

[2] A. Noori, Y. Zhang, N. Nouri and M. Hajivand, "Multi-Objective Optimal Placement and Sizing of Distribution Static Compensator in Radial Distribution Networks With Variable Residential, Commercial and Industrial Demands Considering Reliability," in *IEEE Access*, vol. 9, pp. 46911-46926, 2021.

[3] R. Yan, N. Al-Masood, T. Kumar Saha, F. Bai, and H. Gu, "The anatomy of the 2016 South Australia blackout: A catastrophic event in a high renewable network," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5374–5388, 2018.

[4] O. P. Veloza and F. Santamaria, "Analysis of major blackouts from 2003 to 2015: Classification of incidents and review of main causes," *Electr. J.*, vol. 29, no. 7, pp. 42–49, 2016.

[5] M. M. Aman, G. B. Jasmon, A. H. A. Bakar, H. Mokhlis, and M. Karimi, "Optimum shunt capacitor placement in distribution system - A review and comparative study," *Renew. Sustain. Energy Rev.*, vol. 30, pp. 429–439, 2014.

[6] H. Duran, "Optimum Number, Location, and Size of Shunt Capacitors in Radial Distribution Feeders A Dynamic Programming Approach," *IEEE Trans. Power Appar. and Syst.*, vol. PAS-87, no.9, pp. 1769–1774, 1968.

[7] M. E. Baran and F. F. Wu, "Optimal capacitor placement on radial distribution systems," *IEEE Trans. Power Deliv.*, vol. 4, no. 1, pp. 725–734, 1989.

[8] N. M. Neagle and D. R. Samson, "Loss Reduction from Capacitors Installed on Primary Feeders," *Trans. Am. Inst. Electr. Eng. Part III Power Appar. Syst.*, vol. 75, no. 3, pp. 950–959, 1956.

[9] R. F. Cook, "Analysis of Capacitor Application as Affected by Load Cycle," *Trans. Am. Inst. Electr. Eng. Part III Power Appar. Syst.*, vol. 78, no. 3, pp. 950–956, 1959.

[10] J. V. Schmill, "Optimum Size and Location of Shunt Capacitors on Distribution Feeders," *IEEE Trans. Power Appar. ans Syst.*, vol. 84, no. 9, pp. 825–832, 1965.

[11] K. Mahmoud and M. Lehtonen, "Direct approach for optimal allocation of multiple capacitors in distribution systems using novel analytical closed-form expressions," *Electr. Eng.*, 2020, doi: 10.1007/s00202-020-01073-9.

[12] A. Arief, Antamil, and M. B. Nappu, "An Analytical Method for Optimal Capacitors Placement from the Inversed Reduced Jacobian Matrix," *Energy Procedia*, vol. 100, no. September, pp. 307–310, 2016.

[13] A.Y. Abdelaziz, E.S. Ali, and S.M.A.Elazim, "Flower pollination algorithm and loss sensitivity factors for optimal sizing and placment of capacitors in radial distribution systems," *Int. J. Electr. Power Energy Syst.*, vol. 78, pp. 207-214, 2016.

[14] A. Bayat and A. Bagheri, "Optimal active and reactive power allocation in distribution networks using a novel heuristic approach," *Appl. Energy*, vol. 233–234, pp. 71–85, 2019.

[15] K. Muthukumar and S. Jayalalitha, "Optimal placement and sizing of distributed generators and shunt capacitors for power loss minimization in radial distribution networks using hybrid heuristic search optimization technique," *Int. J. Elect. Power Energy Syst.*, vol. 78, pp. 299–319, 2016.

[16] S. Segura, R. Romero M.J. Rider, "Efficient heuristic algorithm used for optimal capacitor placement in distribution systems," *Int. J. Electr. Power Energy Syst.*, vol. 32, no. 1, pp. 71-78, 2010.

[17] E. A. Almabsout, R. A. El-Sehiemy, O. N. U. An and O. Bayat, "A Hybrid Local Search-Genetic Algorithm for Simultaneous Placement of DG Units and Shunt Capacitors in Radial Distribution Systems," in *IEEE Access*, vol. 8, pp. 54465-54481, 2020.

[18] H.S. Ramadan, A.F. Bendary, and S. Nagy, "Partical swarm optimization for capacitor allocation problem in distribution systems with wind turbine generators," *Int. J. Electr. Power Energy Syst.*, vol. 84, pp. 143-152, 2017.

[19] S. R. Gampa and D. Das, "Optimum placement of shunt capacitors in a radial distribution system for substation power factor improvement using fuzzy GA method," *Int. J. Electr. Power Energy Syst.*, vol. 77, pp. 314–326, 2016.

[20] A. A. Abou El-Ela, R. A. El-Sehiemy, A. M. Kinawy, and M. T. Mouwafi, "Optimal capacitor placement in distribution systems for power loss reduction and voltage profile improvement," *IET Gener. Transm. Distrib.*, vol. 10, no. 5, pp. 1209–1221, 2016.

[21] M. Aryanezhad, "Management and coordination of LTC, SVR, shunt capacitor and energy storage with high PV penetration in power distribution system for voltage regulation and power loss minimization," *Int. J. Electr. Power Energy Syst.*, vol. 100, pp. 178-192, 2018.

[22] M. Resener, S. Haffner, L.A. Pereira, P.M. Pardalos, and M.J.S.Ramos, "A comprehensive MILP model for the expansion planning of power distribution systems – Part I: problem formulation," *Elect. Power syst. Research*, vol. 170, pp. 378-384, 2019.

[23] S. Das, D. Das, and A. Patra, "Operation of distribution network with optimal placement and sizing of dispatchable DGs and shunt capacitors," *Renew. Sustain. Energy Rev.*, vol. 113, 2019, doi: https://doi.org/10.1016/j.rser.2019.06.026.

[24] R.O. Bawazir, and N.S. Cetin, "Comprehensive overview of optimizing PV-DG allocation in power system and solar energy resource potential assessments," *Energy Reorts*, vol. 6, pp. 173-208, 2020.

[25] U. Sultana, A.B. Khairuddin, M.M. Aman, A.S. Mokhtar, and N. Zareen, "A review of optimum DG placement based on minimization of power losses and voltage stability enhancement of distribution system," *Renew. Sustain. Energy Rev.*, vol. 63, pp. 363-378, 2016.

[26] S. Kansal, B. Tyagi, and V. Kumar, "Cost–benefit analysis for optimal distributed generation placement in distribution systems," *Int. J. of Ambient Energy*, 2015, DOI: 10.1080/01430750.2015.1031407.

[27] S. Paul and W. Jewell, "Optimal capacitor placement and sizes for power loss reduction using combined power loss index-loss sensitivity factor and genetic algorithm," *2012 IEEE Power and Energy Society General Meeting*, 2012, pp. 1-8, doi: 10.1109/PESGM.2012.6345635.

[28] M. Chakravorty and D. Das, "Voltage stability analysis of radial distribution networks," *Electr. Power Components Syst.*, vol. 23, pp. 129–135, 2001.

[29] S. Ghosh, V.K. Yadav, and V. Mukherjee, "Evaluation of cumulative impact of partial shading and aerosols on different PV array topologies through combined shannon's entropy and DEA," *Energy*, vol. 144, pp. 765-775, 2018.

[30] J. H. Teng, "A direct approach for distribution system load flow solutions," *IEEE Trans. Power Deliv.*, vol. 18, no. 3, pp. 882–887, 2003.

[31] X. Su et al., "Sequential and Comprehensive BESSs Placement in Unbalanced Active Distribution Networks Considering the Impacts of BESS Dual Attributes on Sensitivity," in *IEEE Transactions on Power Systems*, vol. 36, no. 4, pp. 3453-3464, July 2021.

[32] T. Tewari, A. Mohapatra and S. Anand, "Coordinated Control of OLTC and Energy Storage for Voltage Regulation in Distribution Network With High PV Penetration," in *IEEE Transactions on Sustainable Energy*, vol. 12, no. 1, pp. 262-272, Jan. 2021.

[33] J. Kennedy and E. R., "Particle Swarm Optimization," *Proc. ICNN'95 - Int. Conf. Neural Networks, Perth, Aust.*, vol. 4, pp. 1942–1948, 1995.

[34] C. E. Shannon, "A Mathematical Theory of Communication," *Bell Syst. Tech. J.*, vol. 27, no. 4, pp. 623–656, 1948.

[35] P. Kayal, S. Chanda, and C.K. Chanda, "An analytical approach for allocation and sizing of distributed generations in radial distribution network," *Int. Trans. Elect. Energ. Syst.*, 2017, doi.: https://doi.org/10.1002/etep.2322.

0885-8977 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 09:14:23 UTC from IEEE Xplore. Restrictions apply.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/353326834

Parametric study of solar power tower driven combined supercritical CO 2 cycle and organic Rankine cycle using low GWP fluids

Article · January 2021

DOI: 10.36037/IJREI.2021.5511

CITATIONS
0

READS
17

2 authors, including:



Yunis Khan
Delhi Technological University
10 PUBLICATIONS **19** CITATIONS

SEE PROFILE

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

Thermodynamic analysis of solar basedmodified thermal power plant integrated with Organic Rankine cycle . View project

Thermodynamic analysis of ORC based thermal power plant View project

All content following this page was uploaded by Yunis Khan on 18 July 2021.

The user has requested enhancement of the downloaded file.

International Journal of Research in Engineering and Innovation (IJREI)
journal home page: http://www.ijrei.com
ISSN (Online): 2456-6934



ORIGNAL ARTICLE

Parametric study of solar power tower driven combined supercritical CO_2 cycle and organic Rankine cycle using low GWP fluids

R.S. Mishra, Yunis Khan

Department of Mechanical Engineering, Delhi Technological University, New Delhi, India

Article Information

Received: 28 May 2021
Revised: 9 July July 2021
Accepted: 12 July 2021
Available online: 16 July 2021

Keywords:

Parametric study;
Solar power tower;
Supercritical CO_2 cycle;
Organic Rankine cycle;
Low GWP fluids.

Abstract

In this study, parametric analysis of the solar power tower (SPT) driven combined recuperated supercritical CO_2 Brayton cycle and organic Rankine cycle (ORC) were carried out. Further, low global warming potential and zero ozone depletion fluids were considered in bottoming ORC to reduce harmful effects on environment. A computer program was made in engineering equation solver for simulation. The effects of the SPT design parameters such direct normal irradiation (DNI), concentration ratio (CR), receiver emittance and heat transfer fluid velocity on the system performance were investigated. It was found that thermal performance increased with DNI, CR and heat transfer fluid velocity while decreased with receiver emittance. At 0.95 kW/m^2 of DNI, R1224yd(Z) achieved the highest exergy efficiency, thermal efficiency and net power output of 58.52% and 54.43% and 293.50 kW, respectively. ©2021 ijrei.com. All rights reserved

1. Introduction

Electricity is still primarily produced in the modern world by burning traditional fuels like natural gas, oil, coal, and that not only has a finite lifespan but also emit liquid or gaseous contaminants during service. Solar energy has gained considerable attention as one of the most important candidates for replacing fossil fuels for electricity supply because it is an inexhaustible, renewable, and secure energy source [1, 2]. The initial technology and business planning for CSP (concentrating solar power) technologies such as the dish/engine, power tower, and parabolic trough have recently advanced rapidly around the world. Even then, CSP system's power production efficiencies have been observed as low, which consequently raises the capital expenditure of power production, and further technological innovation of CSP systems must be prioritized. Solar Power Tower (SPT) is the most advanced technology in the CSP. SPT system comprises of a range of sophisticated sub-systems such as a recipient, a tower 75-150 m high and thermal (optional), a heliostat field of 50-150 m^2 and a power conversion system, per heliostat area. Solar radiation focuses on the heliostat receptor field where elevated temperatures heat is produced through a high energy cycle for power production and industrial process supply [3].

1.1 SPT driven cycles

Several SPT driven cycles studies have been performed including combined SCO_2 and trans critical CO_2 cycles [4]; SCO_2 and ORC pre-compression cycles [5]; CO_2 Brayton

Corresponding author: R.S. Mishra
Email Address: hod.mechanical.rsm@dtu.ac.in
https://doi.org/10.36037/IJREI.2021.5511

cycles [6]; SCO_2 recompression with and without main compressor intercooling [7]; three-generation combined cycle [8]; and multi-generation combination cycles [9]. Cycles have been studied for several years. The SCO_2 cycle is a heat-use cycle that may use heat, including geothermal, solar thermal, coal-fueled and natural gas sources from a variety of sources [10]. Various studies were done on solar integrated SCO_2 cycles, such as Singh and Mishra's [11] study on the combined simple recuperated SCO_2 and ORC using parabolic trough solar collectors (PTSC). They discovered that the combined cycle's highest thermal efficiency and exergetic efficiency were found 43.49% and 78.07% at 0.95 kW/m^2 of heat flux, respectively with R407C. In another study, Singh and Mishra [12] investigated improved power generation using PTSC-driven combined recompression of SCO_2 cycle and ORC. At 0.5 kW/m^2 of solar irradiation, the R123-based recompression SCO_2-ORC gave highest exergy efficiency and thermal efficiency of 73.4% and 40.89%, respectively. Singh and Mishra [13] investigated a PTSC-driven simple recuperated SCO_2-VAR cogeneration system for simultaneously cooling and power production. They found that at a DNI of 0.96 kW/m^2 , the highest thermal and exergy efficiency of PTSC were found about 33.9 percent and 65.32%, respectively. On the other hand, exergy destruction rate shows the reverse trend to exergy efficiency, and PTSC have the highest exergy destructions (i.e. 3696 kW). The PTSC operated partial heating SCO_2 ($PSCO_2$) cycle combined with ORC was parametrically analyzed by Khan and Mishra [14]. $PSCO_2$ system was found to be 1-3 percent more thermally efficient than the without partial heating cycle. Also, authors found that incorporating ORC into an existing $PSCO_2$ cycle thermal efficiency increased by 4.47 percent over the standalone $PSCO_2$ cycle. At 950 W/m^2 of solar irradiation, the combined cycle with R1233zd(E) had the highest exergy and thermal efficiency, with 83.26 and 48.61 percent, respectively. Apart from PTSC-driven sCO_2 cycles, a few studies focused on SPT-driven combined SCO_2 cycles have also been carried out, such as The SPT driven combination SCO_2 pre-compression cycle and ORC for waste heat recovery are the results of Khan and Mishra [5] investigation. They found that the use of ORC increases thermal efficiency and net power output of 4.52 and 4.51 percent respectively in the pre-compression cycle. The combined cycle energy, exergy and hottest thermal efficiency of the combination and greatest values of 278.5 kW, 74.06% and 51.83% at 1000 W/m^2 solar radiation with R227ea radiation are enhanced through solar irradiation. The recovery rate of waste heat increases with the efficiency of the heat exchanger. Its greatest value was 0.5673, based on R227ea, at0.95 efficiencies. Khatoon and Kim [4] examined an SPT-driven combined power block system that was incorporated with solar field a thermal energy storage system. The combined system includes SCO_2 recompression Brayton cycle system as a topping cycle and trans critical CO_2 power cycle to recover waste heat. Due to better heat transfer characteristics and temperature glide merit and of the trans critical CO_2 power cycle were chosen over traditional waste heat recovery systems to use other working fluids. Khan and Mishra [1] used ultra-low GWP fluids to conduct a thermo-economic study of the pre-compression SCO_2 cycle with the ORC as bottoming driven by SPT system. Based on the findings, they concluded that HFO performed better than R134a. At 950 W/m^2 of solar irradiation, R1336mzz(Z) had the highest exergy efficiency, thermal efficiency and power output for the combined cycle with 59.60%, 55.02%, and 298.5kW, respectively. The R1336mzz(Z) fluid had the highest waste heat recovery ratio of 0.84, while R134a had just 0.099 at 0.95 heat exchanger effectiveness. Finally, the R1336mzz(Z) and R1234ze(Z) estimated the lowest and highest specific investment costs (SIC), respectively, 2234 and 2290 €/kWe.

Because of its low to medium temperature range, ORC is the best technology for recovering waste heat. Recent research has been done in this area, such as Yagli et al. [15], who used the benefits of the ORC to enhance the net output of a simple gas turbine (GT) situated in a wood production industry. A steam boiler is also connected to the GT in addition to the ORC to boost overall efficiency and generate required steam. Working fluids used in the analysis included toluene, R601, R123, R600, R11 benzene, hexane and cyclohexane. Highest thermal and exergy efficiency of the of the cogeneration system were found 69.19% and 75.51% at ORC turbine inlet parameters. Song et al. [16] proposed a hybrid system for recovering the waste heat from the internal combustion engine (ICE) that integrates the SCO_2 cycle system using the ORC as bottoming cycle from recovering residual heat from the system and from jacket water and exhaust gas. Finally, they estimated that, based on a reference case involving a standalone SCO_2 system for an ICE with 1170 kW of a rated power output, the combined cycle system can produce a maximum net power output of 215 kW at a minimum SIC of 4670 $/kW, which is 58 and 4% more than a standalone SCO_2 power cycles, respectively. It has been found from the literature survey various study were performed on the combined simple SCO_2 Brayton cycle and ORC. However, no study was performed based on solar power tower driven simple recuperated SCO_2 Brayton cycle and ORC using low global warming potential (GWP) and zero ozone depletion potential (ODP) fluids to mitigate worst environmental effects. Therefore, present study deals with the parametric evaluation of the combined cycle. The effects of the SPT design parameters such as direct normal irradiation (DNI), concentration ratio, velocity of heat transfer fluid and receiver emittance on system performance were investigated.

1.2 System description

Current system consists three subsystems first STP system, second simple SCO_2 Brayton cycle and third ORC as bottoming cycle for recovering waste heat. SPT system drives the SCO_2 cycle, through the heat exchanger-1 (HEX1) it supplies heat to the SCO_2 Brayton cycle where power is produced by turbine. After the recuperate still some amount of heat is remaining, through heat exchanger-2 (HEX2) this waste heat recovered by the ORC, where, also some amount of power

is produced by the organic turbine (OT). The whole configuration has been shown in fig. 1.

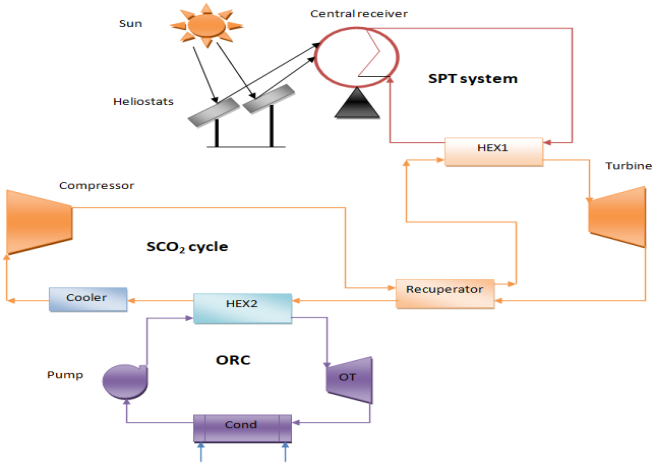


Figure 1. Schematic diagram of SPT driven combined recuperated supercritical CO_2 cycle and ORC

2. Thermodynamic analysis

2.1 Assumptions

SPT powered combined cycle thermodynamic study has been conducted taking account of the following assumptions in support of simulation;

- All components of the system are in steady condition.
- In connecting pipes, friction, pressure and heat loss are overlooked
- Each component's potential and kinetic energy are insignificant
- Input data to facilitate simulation are shown in table 1 with the Heliostat and receiver settings constantly and assumed.
- Molten salt temperature inlet to the HEX1 has been taken 700℃ [17].
- The maximum cycle temperature due to the thermal losses is 50 °C lower than the HEX-inlet salt temperature.
- In the recipients of a mass fraction of 32 percent and 68 percent molded salt HTF was employed in magnesium (MgC_{l2}) and potassium chloride (KCl) mixture [4]. This HTF is the cheapest alternative compared to solar salt and liquid sodium [18] for the heliostat-driven SCO_2-cycle.
- It was also assumed that heliostat receives only beam irradiation.

2.2 Mathematical modeling

This part has established a thermal modelling equation of the proposed system based on the preservation of exergy and energy equations, taking the assumptions stated in the prior section into consideration. Modeling equations for SPT have been taken from the previous studies [1,2,4]. The control volume of each component has also been considered.

Heliostat fields are defined as direct sun heat incidences;

$$\dot{Q}_{solar} = DNI \cdot A_h \cdot N_h \quad (1)$$

Table 1: Input parameters of the proposed model

SPT operational and geometric parameters [1]	
Direct normal irradiation	850 W/m^2 [11]
Temperature of Sun	5800K [19]
Solar multiple	2.8 [4]
Heliostat efficiency	58.71 % [20]
Number of heliostat	141 [21]
Initial temperature difference	15 K [4]
Heliostat's total mirror area	9.04×7.89 [4]
Convective heat loss coefficient	10 W/m^2-K [20]
Solar receiver's temperature approach	423.15 K [20]
Concentration ratio	900 [20]
Convective heat loss factor	1 [20]
Tower height	74.62 m [21]
Absorptance	0.95 [21]
View factor	0.8 [21]
Thermal emittance	0.85 [21]
Combined cycle input data	
SCO_2 turbine inlet pressure	25 MPa [22]
SCO_2 inlet temperature	650 ℃ [4,22]
Compressor inlet temperature	32 ℃ [22, 17]
Compressor inlet pressure	7.5 MPa [22,17]
Compressor isentropic efficiency	0.85 [22,17]
SCO_2 turbine isentropic efficiency	0.88 [22, 17]
SCO_2 topping cycle mass flow rate	1.6 kg/s
Heat exchanger/recuperator effectiveness	0.95 [1,2]
ORC turbine inlet pressure	3 MPa [1,2,11]
Mass flow rate in bottoming ORC	2.7 kg/s
ORC pump's Isentropic efficiency	0.7 [23]
ORC turbine's Isentropic efficiency	0.8 [23]

Where direct normal irradiation (DNI) is DNI, A_h shall be the single area of heliostat (m^2), and N_h shall be the number of the heliostats. However, part of this heat is lost in the environment due to the efficiency of the heliostat. The exact amount of heat generated by the field of heliostat is thus given as;

$$\dot{Q}_h = \dot{Q}_{solar} \cdot \eta_h \quad (2)$$

Where η_h is heliostat's efficiency. The heat is sent in this quantity to the solar receiver where the fluid is transferred. But some of the heat is lost in the atmosphere. Therefore, the heat available at the recipient of the solar center is determined;

$$\dot{Q}_r = \dot{Q}_h \cdot \eta_r = \dot{Q}_h - \dot{Q}_{loss,r} \quad (3)$$

Where, η_r is the receiver thermal efficiency, is defined as;

$$\eta_r = \alpha - \frac{\zeta \cdot f_{view} \cdot \sigma \cdot T_R^4 + h_{conv} \cdot f_{conv} \cdot (T_R - T_{air})}{DNI \cdot \eta_h \cdot CR} \quad (4)$$

Where, T_R is the surface temperature of the receiver and CR

is known concentrated ratio. ζ is the solar emittance. For computing the heat loss, the following expression is used.

$$T_R = T_1 + \delta T_R \quad (5)$$

T_1 is the temperature of the turbine inlet and δT_R is the solar receptor approach temperature.
The solar receiver and the heliostat field are listed in table 1 with operational and geometric specifications.
Moreover, the exergy of any system can be explained as the utmost effort that the system can achieve when the system reaches its dead circumstances. The equation of exergy balance can be established as control volume [24];

$$\sum\left(1-\frac{T_0}{T_J}\right)\dot{Q}_j - \dot{W}_{cv} - \sum(\dot{m}_i E_i) - \sum(\dot{m}_e E_e) - \dot{ED} = 0 \quad (6)$$

Where, $\dot{ED}$ is the rate of exergy degradation, and j relates in a particular State to thermal property. The combined system is determined by the solar exergy inflow by Petela's formula [25];

$$\dot{E}_{solar} = \dot{Q}_{solar}\cdot\left[1+\frac{1}{3}\left(\frac{T_0}{T_{su}}\right)^4 - \frac{4}{3}\left(\frac{T_0}{T_{su}}\right)\right] \quad (7)$$

Where, T_{su} and T_0 are the sun and reference temperature respectively. The useful exergy found by the molten salt in the receiver, is calculated as

$$\dot{E}_r = \dot{m}_{ms}\cdot Cp_{ms}\cdot\left[(T_b - T_a) - \left(T_0\cdot\ln\frac{T_b}{T_a}\right)\right] \quad (8)$$

Assuming throughout constant system chemical exergy and neglecting energy due to velocity and height,the specific physical exergy at j^{th} point is defined as [24];

$$E_j = (h_j - h_0) - T_0(h_j - s_0) \quad (9)$$

Net power output obtained from the combined cycle is defined as;

$$\dot{W}_{net} = \dot{W}_T + \dot{W}_{OT} - \dot{W}_{Comp} - \dot{W}_{pump} \quad (10)$$

Solar powered combined cycle's thermal efficiency is determined as;

$$\eta_{th} = \frac{\dot{W}_{net}}{\dot{Q}_{solar}} \quad (11)$$

The exergy analysis of the combined system using second law of thermodynamics is derived here. The exergy destruction in each component is determined by applying the exergy balance equation (6) for each component after assuming no heat loss in the component [24]. When the exergy destruction rate for each component has been computed, the combined cycle total exergy destruction rate is calculated as;

$$\dot{ED}_{Total} = \dot{ED}_{HEX1} + \dot{ED}_T + \dot{ED}_{OT} + \dot{ED}_{recup} + \dot{ED}_{comp} + \dot{ED}_{cooler} + \dot{ED}_{pump} + \dot{ED}_{HEX2} + \dot{ED}_{cond} + \dot{ED}_{reciever} \quad (12)$$

On the basis of the thermal modeling, numerous mathematical relations are used in the performance analysis of the solar power tower powered combined cycle have been discussed below;
Combined cycle exergy efficiency is determined as [1,2,11,24];

$$\eta_{ex} = 1 - \frac{\dot{ED}_{Total}}{\dot{E}_{solar}} \quad (13)$$

Or

$$\eta_{ex} = \frac{\dot{W}_{net}}{\dot{E}_{solar}} \quad (14)$$

The combined cycle's thermal efficiency can also be defined by the relation between thermal and exergy efficiency of the combined cycle [24];

$$\eta_{th} = \eta_{ex}\cdot\eta_{Carnot} \quad (15)$$

In the SPT powered combined cycle the thermal modeling equations were solved using engineering equation solver (EES) [26].

2.3 Working fluid selection

Working fluids for any device should really be carefully selected because they have an effect on the environment, economic feasibility, and long-term viability. In the receiver, molten salt HTF was made up of a mixture of magnesium dichloride (MgC_{l2}) and potassium chloride (KCl), with mass fractions of 32% and 68%, respectively [4]. This HTF was chosen because, when compared to solar salt and liquid sodium (Na), it is the most cost-effective option for the heliostat-driven sCO_2 cycle [27]. This molten salt's thermo-physical properties in reference Khan and Mishra (2020a). The ORC's working fluid is difficult to select because it loses chemical stability above its optimum temperature, but it achieves the best thermo-physical properties at the right pressure and temperature [18]. To choose suitable fluids for the analysis, various parameters, critical points, such as GWP, thermal stability, and ozone depletion potential (ODP), were analyzed. High GWP fluids, such as hydro fluoro carbons (HCFCs), and high ODP fluids, such as CFCs (chloro fluoro carbons) were replaced by HFO fluids which has ultra-low GWP and ultra-low ODP because GWP is limited at under 150, as restrictions like the European Union [28] prohibit it. Working fluids are called dry, isentropic and moist fluids for the ORC system. Because the expander outlet has high-quality vapour, dry and insulating work is more suited than other fluid. In the present investigation too, the waste heat supply has a low temperature.

This analysis was therefore carried out on the ORC analytics based on nine HFO working fluids such as R1234ze(Z), R1224yd(Z), R1225ye(Z), R1233zd(E), R1234yf, R1243zf, R1234ze(E) and R1336mzz(Z). Flame worthiness, with or without suffix. Normally two classes of toxicity such as lower toxicity (Class A) and higher toxicity (Class B) have been described. There are four flammability classes: 1, 2L, 2 or 3 [29]

2.4 Validation of the proposed cycle

In order to ensure the correct use of the modeling equation, previous studies were used to validate the current cycles. Thermal efficiency has been taken as the validation parameter for both the cycles. Present combined (SCO_2-ORC) cycle was validated with the previous study Singh and Mishra [11] at same baseline conditions respective to references. Thermal efficiency of the cycle was found nearly to the respective previous research as shown in table 2. However, present study was performed with input parameters that are different from the previous studies those were used for the validation purpose.

Table.2. Validation of combined (SCO_2-ORC) cycle

S. No	Baseline conditions.	Thermal efficiency Singh and Mishra [11]	Thermal efficiency Current model	Estimated error
1	Cycle maximum pressure = 25 MPa, Cycle maximum temperature = 400℃, Compressor inlet pressure = 7.4 MPa, Compressor inlet temperature =54.35 ℃, OT inlet pressure=3MPa η_{comp} =0.89, $\eta_{Turbine}$ =0.9 Working fluid in ORC=R245fa	52.03%	52.49%	0.76%

3. Results and discussions

Present study deals with parametric analysis of the SPT driven SCO_2 and ORC system using low GWP fluids. Results were calculated with computational technique considering assumptions and listed data in table 1.

3.1 Influence of the solar irradiation on the system performance

The basic condition for sun irradiation in the Indian climate in Mumbai was 850 W/m2. The impacts of solar irradiation on the efficiency of the system consequently need to be examined, since the current integrated model is powered by a solar power tower. The combined cycle's exergy efficiency was steadily raised with solar irradiation. This is explained by the efficient use of increased solar irradiation on the solar concentrate field. This correlates to an increase in the combined cycle inlet exergy [30].

With solar irradiation, power output and thermal efficiency of the system have also increased. At 950 W/m^2 of solar irradiation, R1224yd(Z) achieved the highest exergy efficiency, thermal efficiency and net power output of 58.52% and 54.43% and 293.50 kW, respectively, as indicated in the Figures 2-4 followed by the fluids R1243zf, R1336mzz(Z), R1233zd(E), R1225ye(Z), R1234ze(Z) and R1234ze(E) and R1234yf. The curve for thermal efficiency and power output has the same pattern as the curve for exergy efficiency. The explanation behind this is that thermal efficiency is directly linked to exergy efficiency [25]. Increase in solar irradiation from 0.4 kW/m^2 to 0.95 kW/m^2, the exergy efficiency, thermal efficiency and power out of the system were increased from 36.73% to 58.52%, 34.16% to 54.42% and 183kW to 293.5kW respectively by R1224ye(Z).

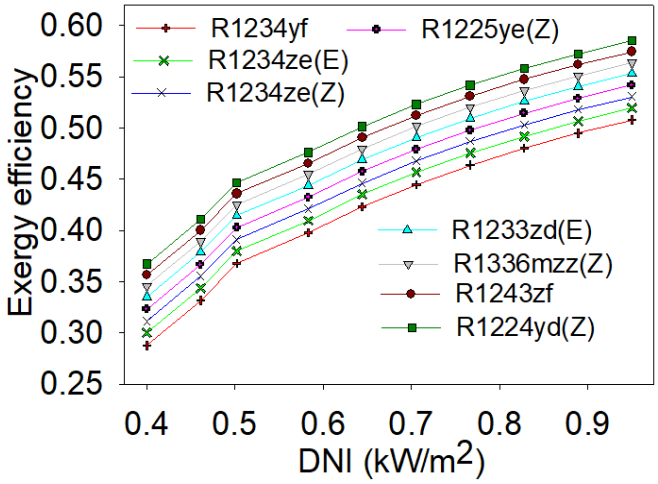


Figure 2. Variation of Exergy efficiency with DNI

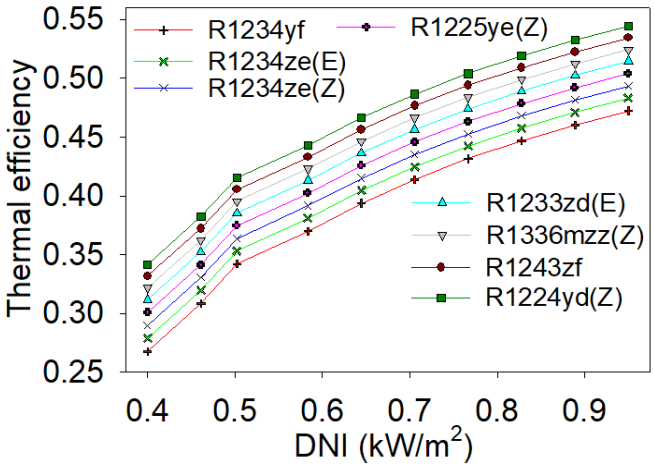


Figure 3. Variation of thermal efficiency with DNI

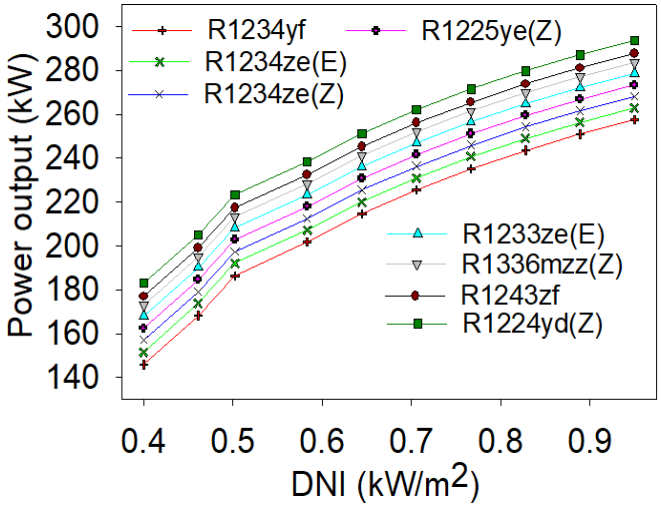


Figure 4. Variation of power output with DNI

3.2 Influence of concentration ratio on the system performance

Another receiver design parameter to consider is the concentration ratio, which has an impact on the combined system's performance. As illustrated in Figures 5-7, increasing the concentration ratio increases combined cycle exergy efficiency and thermal efficiency. As the concentration ratio rises, the receiver efficiency rises, causing the HTF outlet temperature to rise, as the turbine inlet temperature is inversely proportional to the receiver outlet temperature. As a result, as the temperature of the turbine inlet grew, the combined cycle efficiency increased. The fluid R1224yd(Z) achieved the highest exergy efficiency, thermal efficiency and power output once again. Exergy efficiency, thermal efficiency and power output increased from 38.76% to 59.89%, 36.05% to 55.70% and 143.2kW to 342.7kW, respectively, based on R1224yd(Z) when the concentration ratio is increased from 200 to 1400 as shown in figures 5-7.

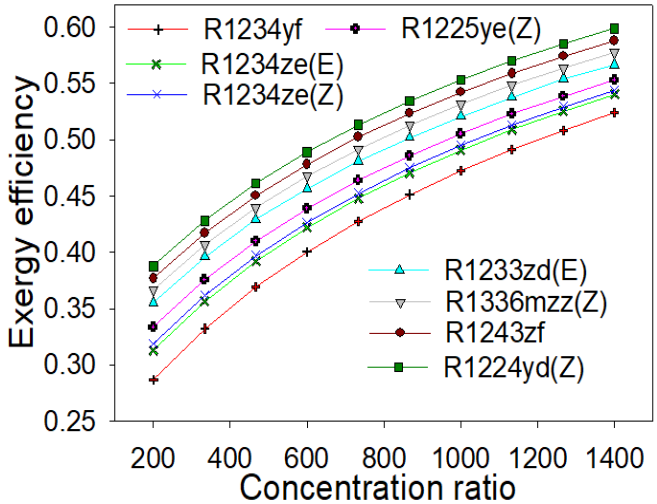


Figure 5. Variation of exergy efficiency with concentration ratio

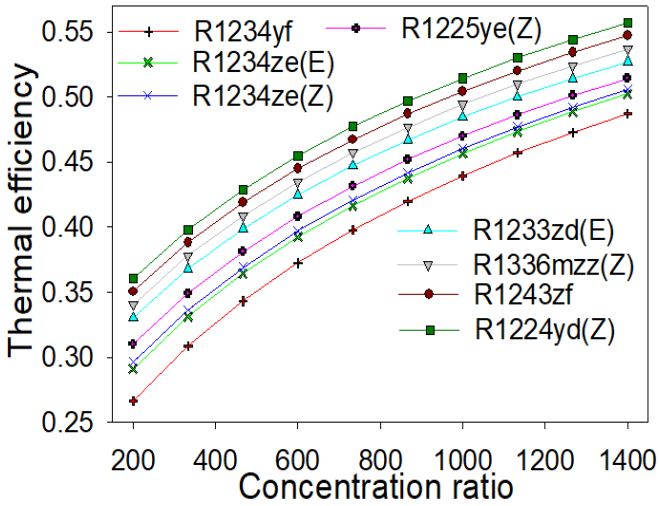


Figure 6. variation of thermal efficiency with concentration ratio

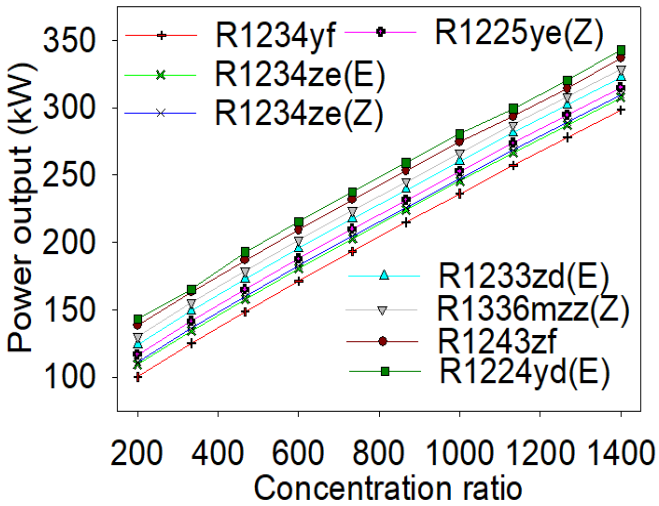


Figure 7. Variation of power output with concentration ratio

3.3 Influence of HTF velocity on system performance

Figures 8-10 show the variation of exergy efficiency, thermal efficiency and power output with HTF velocity in absorber tube. From figure it can be seen that both efficiencies increase with velocity. Reason for increase in second law efficiency with the velocity is that due to increases in velocity of fluid Reynolds number is increased consequently convective heat transfer coefficient increased so much heat is carried with heat transfer fluids so much heat available with HTF. This leads to increase in efficiencies. Highest exergy efficiency, thermal efficiency and power output were obtained for R1224yd(Z) and varies 56.60% to 58.0%, 52.63% to 54.03% and 289.7 to 292.3 kW respectively when velocity varies from 0.01(m/s) to 0.1(m/s) and while lowest values were obtained by the R1234yf among other considered working fluids, it varies from 47.83% to 48.88%, 44.48% to 45.45% and 242.5 to 243.2kW respectively when it velocity varies from 0.01(m/s) to 0.1(m/s).

It was seen that performance improvement slightly varied with the velocity due to effect of standalone cycle only. The performance of bottoming ORC did not affect significantly.

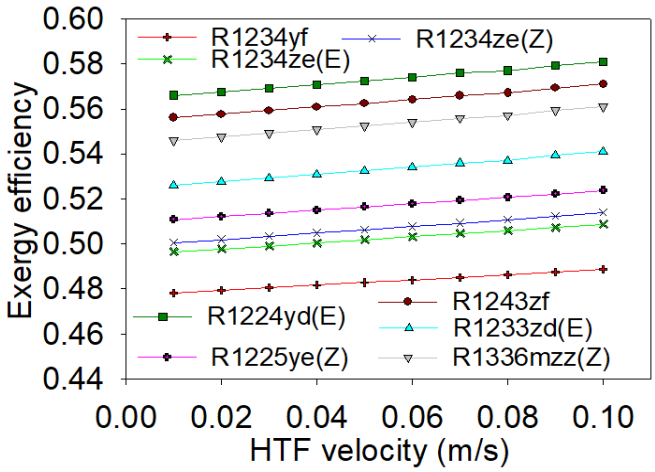


Figure 8. Variation of exergy efficiency with HTF velocity

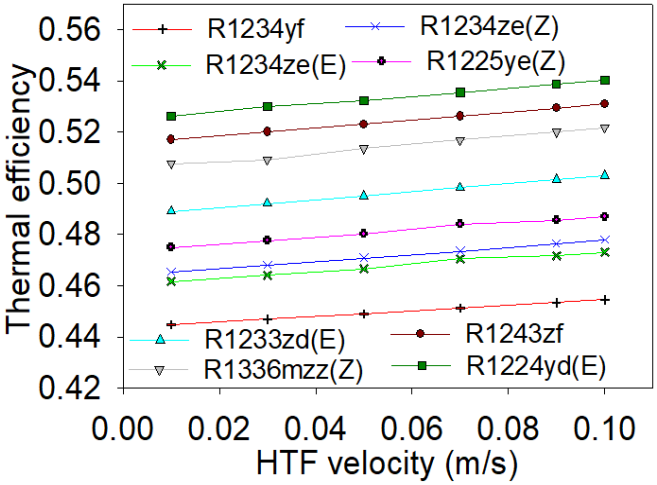


Figure 9. Variation of thermal efficiency with HTF velocity

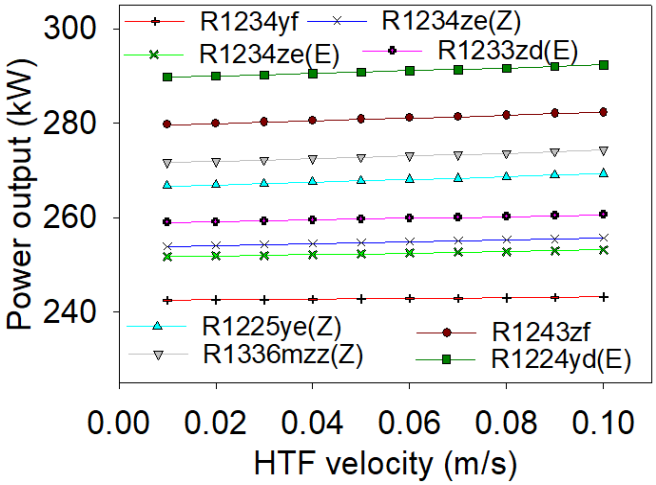


Figure 10. Variation of power output with HTF velocity

3.4 Influence of receiver emittance on system performance

Receiver emittance is the important parameter to be examined because it affects the receiver performance. It is seen in figure 11-13, performance of the combined cycle decreases with the receiver emittance. Receiver's surface temperature is the function of the receiver emittance. Receiver efficiency decreases with the receiver emittance according to Eq. (4). That means more heat loss to the surrounding, consequently less heat energy available to the combined cycle. This leads to decrease in the both efficiencies of the combined cycle. Increase in solar emittance from 0.05 to 0.2 reduces the exergy efficiency, thermal efficiency and output power of the system from 61.22% to 59.66%, 57.29% to 55.84% and 286.4kW to 278.8kW respectively based on the R1224yd(Z) as shown in Figures 11-13. Therefore, it becomes necessary to decrease the solar emittance while designing the SPT to get better performance of combined cycle for power generation.

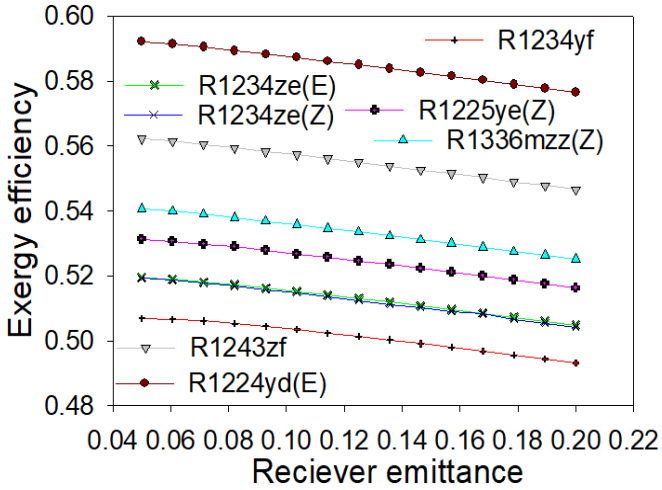


Figure 11. Variation of exergy efficiency with receiver emittance

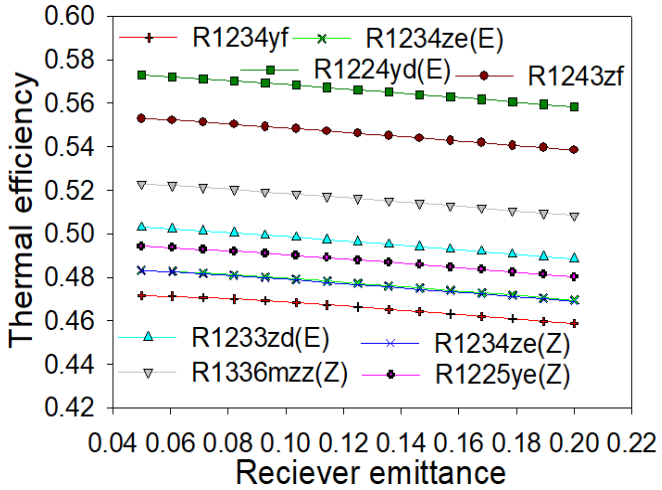


Figure 12. Variation of thermal efficiency with receiver emittance

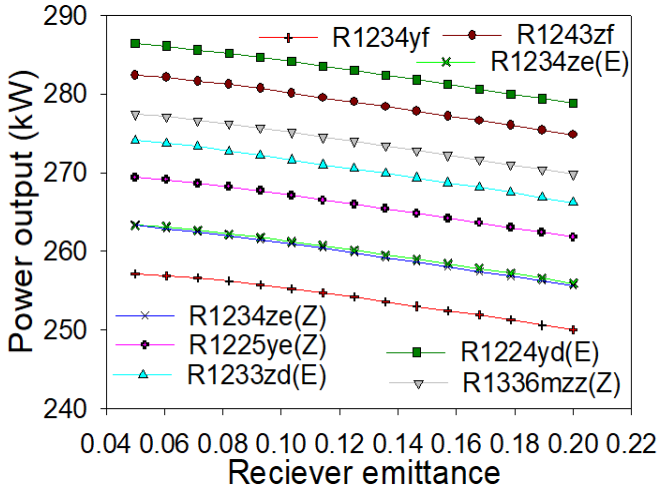


Figure 13. Variation of power output with receiver emittance

Conclusions

- Performance of the system increased with the DNI. Maximum performance of combined cycle was found with R1224yd(E) fluid followed by R1243zf, R1336mzz(Z), R1233zd(E), R1225ye(Z), R1234ze(Z), R1234ze(E) and R1234yf at present input conditions. Maximum exergy efficiency, thermal efficiency and output power were increased from 36.73% to 58.52%, 34.16% to 54.42% and 183kW to 293.5kW respectively when DNI increased from 0.4 kW/m^2 to 0.95 kW/m^2 based on R1224yd(Z) fluid.
- Increase in concentration ratio from 200 to 1400 increases the highest thermal and exergy efficiency and output power increased from 38.76% to 59.89%, 36.05% to 55.70% and 143.2kW to 342.7kW with fluid R1224yd(Z).
- Performance of combined cycle increased with velocity of HTF in receiver. Highest exergy efficiency, thermal efficiency and power output were obtained for R1224yd(Z) and varies 56.60% to 58.1%, 52.92% to 54.4% and 289.7 to 292.3kW respectively when velocity varies from 0.01(m/s) to 0.1(m/s).
- Apart from this R1224yd(Z) may be recommended for better performance of the combined cycle based on current input conditions.
- Current study limited to the parametric analysis and effects evaluation of few selected SPT design parameters on combined cycle. Further, this system can be analyzed with more SPT design parameters.

References

[1] Khan Y, Mishra RS. Thermo-economic analysis of the combined solar based pre-compression supercritical CO_2 cycle and organic Rankine cycle using ultra low GWP fluids. Thermal Science and Engineering Progress 2021; 23:100925.doi.10.1016/j.tsep.2021.100925.

[2] Khan Y, Mishra RS. Performance analysis of solar driven combined recompression main compressor intercooling supercritical CO_2 cycle and organic Rankine cycle using low GWP fluids. Energy and Built Environment 2021.doi.10.1016/j.enbenv.2021.05.004.

[3] Farges O, Bézian J, El-Hafi M. Global optimization of solar power tower systems using a Monte Carlo algorithm: Application to a redesign of the PS10 solar thermal power plant. Renewable Energy 2018;119:345-53.

[4] Khatoon S, Kim M. Performance analysis of carbon dioxide based combined power cycle for concentrating solar power. Energy Conversion and Management 2020.doi: 10.1016/j.enconman.2019.112416.

[5] Khan Y, Mishra RS. Performance evaluation of solar-based combined pre-compression supercritical CO_2 cycle and organic Rankine cycle. International journal of Green energy 2020. doi:10.1080/15435075.2020.1847115.

[6] Alsagri AS, Chiasson A, Gadalla M. Viability Assessment of a Concentrated Solar Power Tower with a Supercritical CO2 Brayton Cycle Power Plant. Journal of Solar Energy Engineering 2018. doi:10.1115/1.4043515

[7] Ma Y, Morozyuk T, Liu M, Yan J, Liu J. Optimal integration of recompression supercritical CO_2 Brayton cycle with main compression intercooling in solar power tower system based on exergoeconomic approach. Applied Energy 2019; 242:1134-54.

[8] Shukla AK., Sharma A, Sharma M, Nandan G.Thermodynamic investigation of solar energy-based triple combined power cycle, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects 2018. doi: 10.1080/15567036.2018.1544995.

[9] Abid M, Adebayo VO, Atikol U. Energetic and exegetic analysis of a novel multi-generation system using solar power tower. International journal of exergy 2019.doi:10.1504/IJEX.2019.100364

[10] Ahn Y, Bae SJ, Kim M, Cho SK, Baik S, Lee JI, Cha JE. Review of supercritical CO_2 power cycle technology and current status of research and development, Nuclear Engineering Technology 2015;47: 647–61.

[11] Singh, H., and R. S. Mishra. 2018. Performance analysis of solar parabolic trough collectors driven combined supercritical CO2 and organic Rankine cycle. Engineering Science and Technology, an International Journal 21:451–44. doi:10.1016/j.jestch.2018.03.015.

[12] Singh, H. and R.S. Mishra. 2018. Energy and exergy-based performance evaluation of solar powered combined cycle (recompression supercritical carbon dioxide cycle/organic Rankine cycle). Clean Energy, 2018, 1–14.

[13] Mishra, R.S and H. Singh. 2018. Detailed parametric analysis of solar driven supercritical CO_2 based combined cycle for power generation, cooling and heating effect by vapor absorption refrigeration as a bottoming cycle. Thermal Science and Engineering Progress 8:397–410.

[14] Khan Y, Mishra RS. Parametric (exergy-energy) analysis of parabolic trough solar collector-driven combined partial heating supercritical CO_2 cycle and organic Rankine cycle. Energy sources, part a: recovery, utilization, and environmental effects 2020.doi:10.1080/155

[15] HüseyinYağlı, YıldızKoç, and Hüseyin Kalay. 2021. Optimisation and exergy analysis of an organic Rankine cycle (ORC) used as a bottoming cycle in a cogeneration system producing steam and power Sustainable Energy Technologies and Assessments. 44:100985.

[16] Jian Song , Xiaoya Li , Kai Wang , Christos N. Markides. 2020. Parametric optimization of a combined supercritical CO2 (S-CO2) cycle and organic Rankine cycle (ORC) system for internal combustion engine (ICE) waste-heat recovery. Energy Conversion and Management, 218:112999.

[17] Reyes-Belmonte, M.A., A. Sebastián, and M. Romero.2016.Optimization of a recompression supercritical carbon dioxide cycle for an innovative central receiver solar power plant. Energy.112 :17–27.

[18] Polimeni, S., M. Binotti, L. Moretti, and G. Manzolini. 2018. Comparison of sodium and KCl-$MgCl_2$ as heat transfer fluids in CSP solar tower with SCO_2 power cycles. Solar Energy.162:51024.

[19] Bejan A, Kearney DW, Kreith F. Second law analysis and synthesis of solar collector systems. J. Sol. Energy Eng. Trans. ASME 1981;103:23–28.

[20] Ho CK., Iverson BD. Review of high-temperature central receiver designs for concentrating solar power. Renewable Sustainable Energy Reviews 2014;29:835–46.

[21] Wang X, Liu Q, Lei J, Han W, Jin H. Investigation of thermodynamic performances for two-stage recompression supercritical CO2 Brayton

cycle with high temperature thermal energy storage system. Energy Conversion Management 2018;165:477–87.
[22] Neises T, Turchi C. A Comparison of Supercritical Carbon Dioxide Power Cycle Configurations with an Emphasis on CSP Applications. Energy Procedia 2014;49:1187–96.
[23] Song J, Li X, Ren X, Gu C. Performance analysis and parametric optimization of supercritical carbon dioxide (S-CO2) cycle with bottoming Organic Rankine Cycle (ORC). Energy 2018;143:406-16.
[24] Cengel YA, Boles MA. Thermodynamics An Engineering Approach. 5th edition;McGraw-Hill publication. New York;2004.
[25] R. Petela, Exergy analysis of the solar cylindrical-parabolic cooker, Sol. Energy 79 (2005) 221–233.
[26] Klein SA. Engineering Equation Solver (EES), Academic Commercial V7.714. F-Chart Software 2020, www.fChart.com.
[27] Koc,Y., H. Yagli, and A. Koc. 2019. Exergy analysis and performance improvement of a subcritical/supercritical Organic Rankine Cycle (ORC) for exhaust gas waste heat recovery in a biogas fuelled Combined Heat and Power (CHP) 520 engine through the use of regeneration. Energies 12 (4):575. doi:10.3390/en12040575.
[28] Moloney, F., E. Almatrafi, D.Y. Goswami. 2017.Working fluids parametric analysis for the regenerative supercritical organic Rankine cycle for medium geothermal reservoir temperatures. Energy Procedia.129:599- 606.
[29] Calm,J.M.(1994):Referegerant safety. ASHRAE Journal.36.(7):17-26.
[30] Al-Sulaiman FA. Exergy analysis of parabolic trough solar collectors integrated with combined steam and organic Rankine cycles. Energy Conversion and Management 2014;77:441–49.
[31] Khan Y and Mishra RS. 2021. Performance comparison of basic and parallel double evaporator organic rankine cycle integrated with solar based supercritical CO_2 cycle. Journal of Thermal Engineering. (accepted)

Nomenclature

$\dot{E}_{solar}$	Solar exergy (kW)
$\dot{Q}_h$	Actual solar heat received by heliostat field (kW)
$\dot{Q}_{loss,r}$	Heat loss from the receiver (kW)
$\dot{Q}_r$	Heat received by central receiver (kW)
$\dot{Q}_{solar}$	Solar heat received by heliostat field (kW)
A_h	Heliostat area (m^2)
$\dot{E}$	Exergy rate (kW)
$\dot{E}D$	Rate of exergy destruction (kW)
f_{view}	Receiver's view factor
h_{conv}	Coefficient for convection heat loss (W/ m^2-K)
$\dot{m}$	Mass flow rate (kg/s)
N_h	Number of heliostat
$\dot{Q}$	Heat rate in (kW)
T_R	Receiver surface temperature (K)
$\dot{W}$	Power (kW)
η_{ex}	Exergy efficiency
η_h	Heliostat efficiency
η_r	Receiver thermal efficiency
η_{th}	Thermal efficiency
Cp	Specific heat (kJ/kg-K)
DNI	Direct normal irradiation (W/m^2)
h	specific enthalpy (kJ/kg)
s	specific entropy (kJ/kg-K)
SCO_2	Supercritical carbon dioxide
T	Temperature (K)

Abbreviations

Comp	Compressor
Cond	condenser
CR	concentration ratio
CV	Control volume
HEX1	heat exchanger-1
HEX2	heat exchanger-2
ORC	Organic Rankine cycle
OT	ORC turbine
recup	Low temperature recuperator
SPT	solar power tower

Subscripts

e	exit
0	dead condition
r	receiver
h	heliostat
i	inlet
j	particular state
su	Sun
ms	molten salt

Greek letters

η	Efficiency
$\boldsymbol{\varepsilon}$	Effectiveness
α	Solar absorbance
δ	Change in property
β	Sun's subtended cone half angle(rad)
σ	Stephen Boltzmann constant (W/m)

Cite this article as: R.S. Mishra, Yunis Khan, Parametric study of solar power tower driven combined supercritical CO2 cycle and organic Rankine cycle using low GWP fluids, International journal of research in engineering and innovation (IJREI), vol 5, issue 5 (2021), 291-299. https://doi.org/10.36037/IJREI.2021.5511.

View publication stats

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

Performance Analysis of Optical Communication System for Various Filtering Operations Using MIMO FSO System under Fog Condition

Anurag Chauhan
dept. of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
anuragchauhan@dtu.ac.in

Archana Kumari
dept. of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
archanakumari_2k17ec35@dtu.ac.in

Priyanka Meena
dept.of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
priyankameena_2k17ec132@dtu.ac.in

Sanjay Kumar Meena
dept. of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
sanjaykumarmeena_2k17ec148@dtu.ac.in

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498509

***Abstract*— In this paper FSO MIMO technology has been used to analyze distinct pre and pre-post amplified optical filters under fog condition. Optical filters provide larger bandwidth with MIMO technology for several applications in 5G communication. Gaussian, Bessel and Rectangle optical filters have been used in this analysis for 5G optical communication. Optical filters are compared in terms of different parameters i.e. beam divergence, range, diameter of transmitter and receiver aperture under fog condition. The results have been plotted for various optical filters based on different parameters. The Q-factor is 765.80 at the range of 100 m for Gaussian optical filter.**

Keywords—FSO, MIMO, 5G communication, Q-factor, Optical amplified filter.

I. INTRODUCTION

Free Space Optics (FSO) is now being considered as one of the most useful technology in terms of wireless communication. FSO provides high data capacity and can transfer a large amount of information. In several areas optical fiber cable does not work properly but with the help of FSO communication, this problem can be solved. FSO is the main reason for the decrease in demand for optical fiber cables. FSO has several benefits such as no connection requirements between transmitter and receiver and has an easy installation process [1]. The FSO channel relies on the inside elements and outside components of the framework. The inside factors incorporates wavelength, power, transfer speed, data rate and receiver sensitivity while the outside factors incorporate attenuation, window misfortune and scintillations [2,3]. FSO also provides less power consumption in comparison to optical fiber cable [4]. Advancement in Multiple Input Multiple Output (MIMO) technology shows that the communication can also take place at larger link along with better efficiency [5].

Recently 5G based communication systems are being implemented with the help of optical wireless channels that has been wide unfold in several techniques like millimeter (mm) signal 5G communication, cellular construction 5G and 5G backhaul wireless link [3-8]. Recently, FSO based system has been implemented using distinct optical filters in 5G communication. In this reported work, the MIMO FSO based on 5G optical communication system is designed and system performance is analyzed for parameters such as beam divergence, range, transmitter and receiver aperture respectively [6].

II. SYSTEM DESCRIPTION

The block diagram represented in Fig.1 shows the FSO based MIMO when no optical filter is used. The block diagram shows the transmitter part which contains Continuous wave laser. Mach-Zehnder modulator has been used for modulation, Pseudo Random Bit Sequence generator at the data rate of 20 Gbps, and non-return to zero (NRZ) pulse generator has been used for encoding method [9]. There is a need of an optical amplifier with optical gain of 20 dB to improve the strength for transmission of signal in short range communication [6,10]. Hence, an optical amplifier has been placed before the FSO channel with optical gain of 20 dB. The signal is passed to the receiver side after amplification through FSO link under fog attenuation. The receiver section includes Avalanche Photodiode Detector (APD) to detect the optical signal and BER analyzer to analyze Q-factor.

The block diagram shown in Fig. 2, represents the performance of the modified model with pre-amplified optical filter [11]. Bessel, Gaussian and Rectangle optical filters are studied. Optical amplifier is used for pre-amplification for short range communication set at 20 dB [6]. MIMO FSO channel is used to attain minimum BER as well as obtain high quality factor with the help of BER analyzer. The modified model consists of pre-amplified optical filter which helps improves 5G communication [6,15].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:54:38 UTC from IEEE Xplore. Restrictions apply.

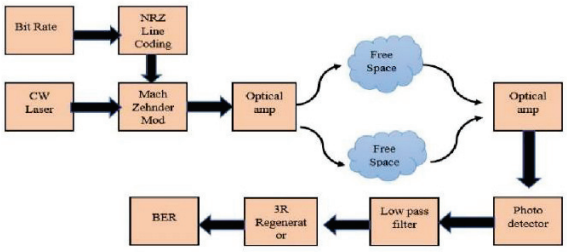


Fig. 1. The MIMO FSO model without using optical filter (previous model)

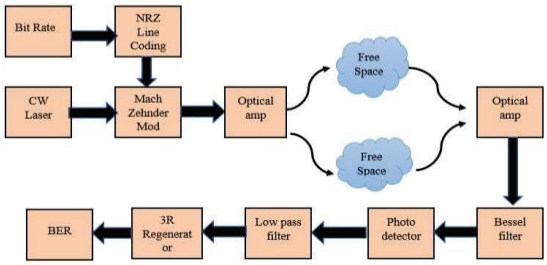


Fig. 2. Modified model using pre-amplified optical filter.

The block diagram shown in Fig 3, represents the performance of the pre-post amplified filtering that improves the Q-factor. Optical amplifier can be used for pre-post amplification set at 20 dB for small range communication. [6]. It is observed that there is improvement in the system model when transmitter aperture is decreased and the receiver aperture is increased [12].

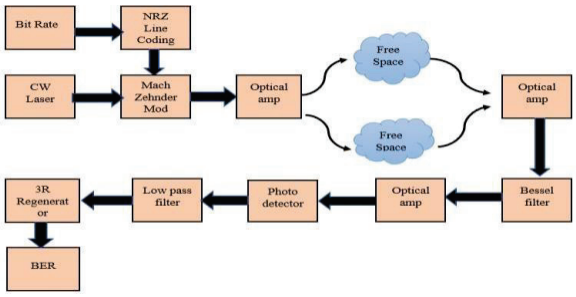


Fig. 3. Proposed model using pre-post amplified optical filter

III. Schematic Design

The schematic diagram for the system model is shown in Fig.4 which consists of transmitter section, 2× 2 MIMO FSO channels and receiver section without including optical filters. Fig.5 and Fig.6 shows the system model which consists of pre and pre-post amplifier optical filters respectively for obtaining the high bandwidth required for 5G application. The system performance is analyzed under fog condition. Input power is taken to be 0 dBm and range, beam divergence, transmitter and receiver aperture are varying for all three optical filters. Range has been varied from 100-1000 m, beam divergence from 2 mrad to 10 mrad [6], transmitter aperture from 140-26 cm and receiver aperture from 140-210 cm [14]. Attenuation for fog condition is taken to be 26 dB/km due to more turbulence effect of fog on FSO link [13].

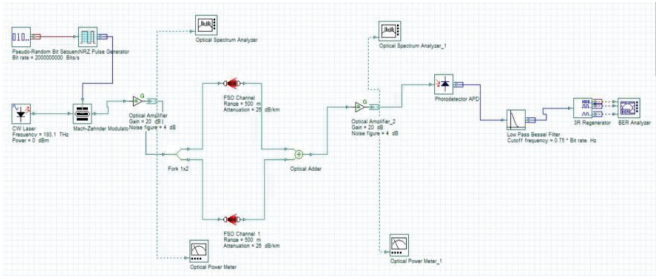


Fig. 4. Schematic design for previous model

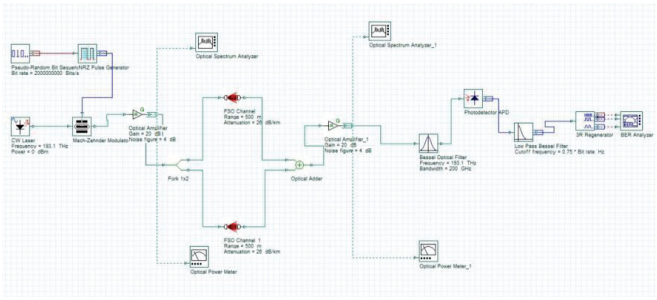


Fig. 5. Schematic design for modified model

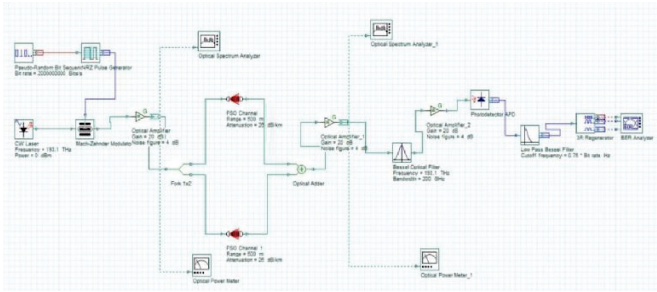


Fig. 6. Schematic design for proposed model

IV. Result and Disscusion

The modified model provides Q-factor for the range below 500m as compared to previous model (without using optical filter). The Q-factor achieved in case of previous model is less than modified model with respect to varying range. Q-factor decreases with increasing link distance as shown in Fig.7, Fig.8 and Fig.9 using Gaussian, Bessel and Rectangle optical filter.

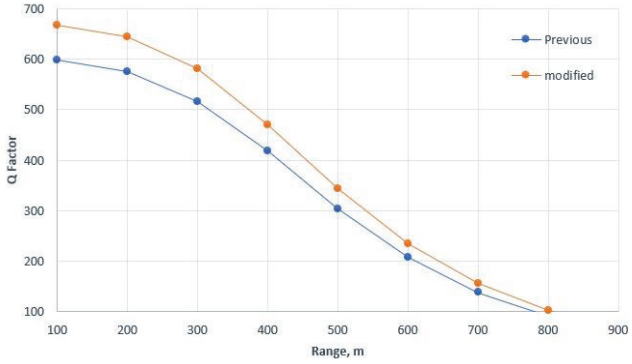


Fig. 7. Q-factor vs. Range for previous and modified model with Bessel filter.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:54:38 UTC from IEEE Xplore. Restrictions apply.

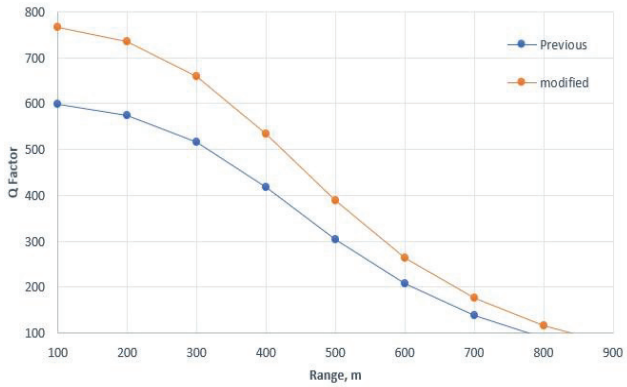


Fig. 8. Q-factor vs. Range for previous and modified model with Gaussian filter.

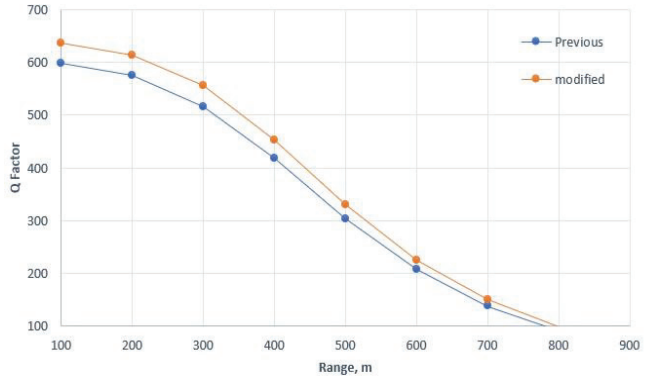


Fig. 9. Q-factor vs. Range for previous and modified model with Rectangle filter.

The graphs shown in Fig.10 present the analysis between Q-factor vs. range for pre-amplified optical filters i.e. Bessel, Gaussian and Rectangle optical filter. The Q-factor was maximum for range 100 m in case of Gaussian optical filter in comparison to both Bessel and Rectangle optical filter. The Q-factor starts deteriorating beyond the range of 1 km for Bessel, Gaussian and Rectangle filters.

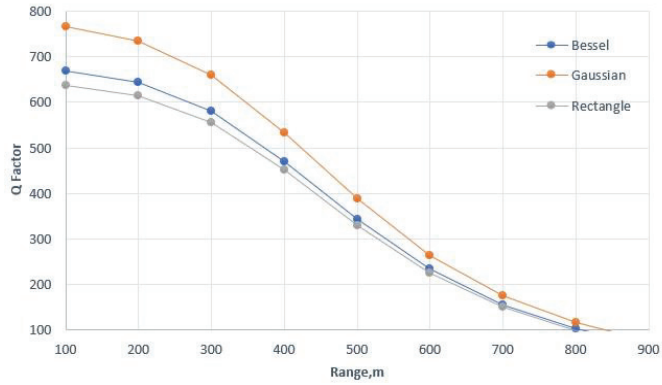


Fig. 10. Q-factor vs. Range of pre-amplified optical filters

The graph shown in Fig.11, presents the analysis between Q-factor and beam divergence respectively, it can be observed that the Q-factor value is highest when beam divergence is below 4 mrad for pre-amplified optical filters. The Q-factor starts deteriorating when beam divergence is increased. The highest Q-factor is achieved in case of Gaussian filter as compared to Bessel and rectangular optical filter.

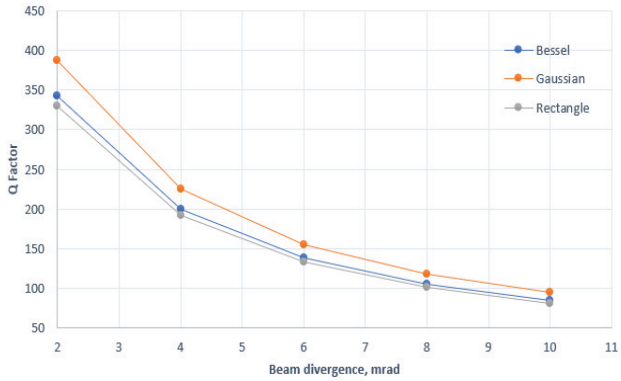


Fig. 11. Q-factor vs. beam divergence of pre- amplified optical filters

The graph shown in Fig.12 and Fig.13, represents the analysis of Q-factor vs diameter of transmitter and receiver aperture respectively. Q-factor value is increases when diameter of transmitter aperture decreases and diameter of receiver aperture increases in case of pre-post amplified optical filter [14]. The Q-factor gives highest value in case of Gaussian as compared to Bessel and Rectangular optical filter.

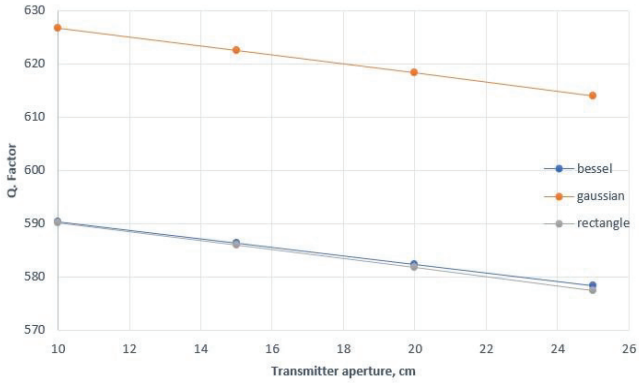


Fig. 12. Q-factor vs. transmitter aperture diameter

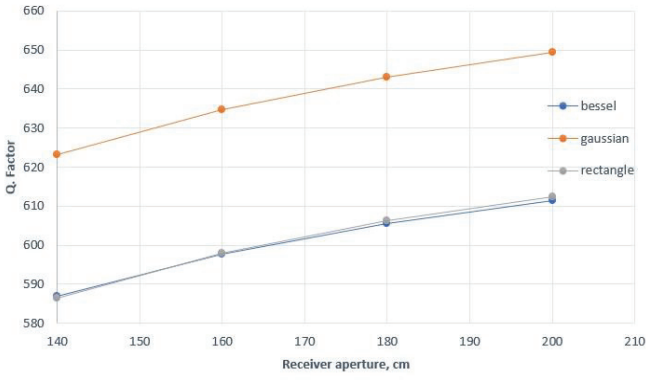


Fig. 13. Q-factor vs. Receiver aperture diameter

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:54:38 UTC from IEEE Xplore. Restrictions apply.

The following tables given below shows the comparative analysis of different pre and pre-post amplified optical filters for different parameters such as beam divergence, range, diameter of transmitter and receiver aperture.

TABLE I. THE GIVEN TABLE REPRESENTS THE Q-FACTOR ANALYSIS OF PREVIOUS AND MODIFIED MODEL WITH VARIOUS FILTERS FOR VARYING RANGE.

Range (m)	Previous Model	Modified Model (Bessel)	Modified Model (Gaussian)	Modified Model (Rectangle)
		Q-factor		
100	598.85	668.01	765.80	635.96
200	574.82	643.47	734.99	614.24
300	516.20	580.08	659.69	555.58
400	417.45	470.31	532.88	451.85
500	304.24	342.91	387.79	330.17
600	207.64	234.16	264.48	225.61
700	137.60	155.70	175.59	149.94
800	90.19	102.83	115.73	98.92
900	58.98	67.90	76.24	65.25
1000	38.29	44.83	50.51	43.37

TABLE II. THE GIVEN TABLE REPRESENTS THE Q-FACTOR ANALYSIS OF PREVIOUS AND MODIFIED MODEL WITH VARIOUS FILTERS FOR VARYING BEAM DIVERGENCE.

Beam Div. (mrad)	Previous Model	Modified Model (Bessel)	Modified Model (Gaussian)	Modified Model (Rectangle)
		Q-factor		
2	598.85	342.91	387.79	330.17
4	574.82	199.09	224.74	191.80
6	516.20	137.75	155.25	132.61
8	417.45	104.65	117.80	100.68
10	304.24	84.08	94.53	80.82

TABLE III. THE GIVEN TABLE REPRESENTS THE Q-FACTOR ANALYSIS OF PROPOSED MODEL WITH VARIOUS FILTERS FOR VARYING DIAMETER OF TRANSMITTER APERTURE.

Tx. aperture (cm)	Rx. aperture (cm)	Proposed Model (Bessel)	Proposed Model (Gaussian)	Proposed Model (Rectangle)
		Q-factor		
25	80	578.26	613.91	577.39
20	80	582.34	618.27	581.69
15	80	586.33	622.54	585.91
10	80	590.24	626.72	590.03

TABLE IV. THE GIVEN TABLE REPRESENTS THE Q-FACTOR ANALYSIS OF PROPOSED MODEL WITH VARIOUS FILTERS FOR VARYING DIAMETER OF RECEIVER APERTURE.

Tx. aperture (cm)	Rx. aperture (cm)	Proposed Model (Bessel)	Proposed Model (Gaussian)	Proposed Model (Rectangle)
		Q-factor		
100	140	586.89	623.14	586.50
100	160	597.74	634.76	597.98
100	180	605.55	643.13	606.24
100	200	611.33	649.33	612.37

TABLE V. THE GIVEN TABLE REPRESENTS THE Q-FACTOR ANALYSIS OF PROPOSED MODEL WITH VARIOUS FILTERS FOR VARYING DIAMETER OF TRANSMITTER AND RECEIVER APERTURE.

Tx. aperture (cm)	Rx. aperture (cm)	Proposed Model (Bessel)	Proposed Model (Gaussian)	Proposed Model (Rectangle)
		Q-factor		
290	240	574.11	609.47	573.01
300	250	574.94	611.25	577.56
310	260	577.31	612.84	579.05

V. CONCLUSION

On the basis of performance analysis it can be concluded that the modified and proposed model is better than the previous model. The result of simulation graphs between Q-factor versus range and Q-factor versus beam divergence proves that modified model is better with Gaussian optical filter as compared to Bessel and Rectangle optical filter. In proposed model Q-factor achieved is maximum in case of Gaussian optical filter when transmitter aperture is minimum and receiver aperture is maximum. This Analysis shows that Q-factor depends upon beam divergence, link distance, and diameter of transmitter aperture and receiver aperture. The modified and proposed model Q-factor represents the improvement of the system. Different optical filters using pre-post amplification provides high bandwidth for 5G application.

REFERENCES

[1] Roy R, Babu JS, "Simulation and performance analysis of free space optical systems using multiple TX/RX and polarized COOFDM techniques under atmospheric disturbances," Int J Eng Res: Gen Sci 2015;3(1):743–9.

[2] S. Robinson, S. Jasmine, R. Pavithra, "Investigation on Hybrid WDM (DWDM+CWDM) Free space optical communication," Journal of Communication Technology, vol. 6, No. 4, pp. 11871192, 2015.

[3] P. Sathya and S. Robinson, "Impact of Modulation Techniques for Hybrid-WDM Based FSO," 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT), Coimbatore, India, 2018, pp. 1-6, doi: 10.1109/ICCTCT.2018.8550907.

[4] Ali, H. A. E. M., Said, E.-S. S. A., Yousef, M. E., "Effect of Environmental Parameters on the Performance of Optical Wireless Communications," International Journal of Optics, 2019, 1–12. doi: https://doi.org/10.1155/2019/1828275.

[5] M. Matalatala, M. Deruyck, E. Tanghe, L. Martens, and W. Joseph, "Performance evaluation of 5G millimeter-wave cellular access networks using a capacity-based network deployment tool," Mobile Information Systems, vol. 2017, Article ID 3406074, 21 pages, 2017.

[6] K.H. Shakthi Murugan, M.Sumathi, "Design and Analysis of 5G Optical Communication System Various Filtering Operations using Wireless Optical Transmission," Results in Physics, Volume 12, March 2019, Pages 460-468.

[7] C. Wang et al., "Cellular architecture and key technologies for 5G wireless communication networks," in IEEE Communications Magazine, vol. 52, no. 2, pp. 122-130, February 2014, doi: 10.1109/MCOM.2014.6736752.

[8] Jan Bohata, Matěj Komanec, Jan Spáčil, Zabih Ghassemlooy, Stanislav Zvánovec, and Radan Slavík, "24–26 GHz radio-over-fiber and free-space optics for fifth-generation systems," Opt. Lett. 43, 1035-1038 (2018).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:54:38 UTC from IEEE Xplore. Restrictions apply.

[9] Md Noor, N. H., Naji, A. W., & Al-Khateeb, W, "Performance Analysis of a Free Space Optics Link with Multiple Transmitters/Receivers," IIUM Engineering Journal, 13(1). https://doi.org/10.31436/iiumej.v13i1.271.

[10] Jay arajana P, Kuppusamyb PG, Sundararajan TVP, Thiyagupriyadharsana MR, AhamedYasara Z, Maheswara R, Amiri Iraj S, "Analysis of temperature based power spectrum in EDFA and YDFA with different pump power for THz applications," Results Phys 2018;10:160–3.

[11] Shakthi Murugan K. H. and Sumathi M, "BER and Eye Pattern Analysis of 5G Optical Communication System withFilters," 2019 IEEE International Conference on Electrical, Computer and Communication Technologies 1-6.

[12] H. Kaur, A. Bhardwaj and G. Soni, "Performance improvement of free space optical link by varying transmitter and receiver aperture diameter," 2015 Global Conference on Communication Technologies (GCCT), Thuckalay, India, 2015, pp. 761-766, doi: 10.1109/GCCT.2015.7342766.

[13] Shakir, F., Ali, M. A. A., & Ameer, F, "Utilization of MIMO Concept for Optical Communication System under Fog Condition," ECTI Transactions on Electrical Engineering, Electronics, and Communications, 17(2), 130-135. https://doi.org/10.37936/ecti-eec.2019172.219183.

[14] M. Nandi, E. Roy, S. Chakraborty and T. Basu, "Study of Variation of Q-Factor Using Minimum Power Under Different Weather Conditions," 2019 International Conference on Opto-Electronics and Applied Optics (Optronix), Kolkata, India, 2019, pp. 1-6, doi: 10.1109/OPTRONIX.2019.886239.

[15] Optiwave Corporation, OptiSystem Overview, https://optiwave.com/resources/applications-resources/opticalcommunication-system-design/.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:54:38 UTC from IEEE Xplore. Restrictions apply.

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

Performance Estimation of Power Factor Corrected Converters for BLDC Motor Drive

Utkarsh Bhardwaj
Electrical Engineering Department
Delhi Technological University
Delhi, India
utkarshbhardwaj.ee@gmail.com

Abhishek Verma
Electrical Engineering Department
Delhi Technological University
Delhi, India
vermabhishek2000@gmail.com

Sanjay Kumar
Electrical Engineering Department
Delhi Technological University
Delhi, India
sanjuimagination2000@gmail.com

Saurabh Mishra
Electrical Engineering Department
Delhi Technological University
Delhi, India
saurabh.mishra1410@gmail.com

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498514

***Abstract*— A comparative analysis of three Power Factor Correction (PFC) converters i.e. Single-Ended Primary Inductance Converter (SEPIC), Cuk Converter and Zeta Converter have been done on the BLDC motors. BLDC motors are generally used for low power application thus they are powered through a 1-Φ sinusoidal AC supply with a Full Bridge Diode Rectifier (DBR). The DBR decreases the power factor and thereby the power quality of the complete system and because of this various power quality concerns have been raised in BLDC motors. A PFC converter offers a good solution in meeting the power quality demands and increasing the reliability of the system. For getting an output voltage for a corresponding speed at the end of the PFC converter, a voltage follower scheme has been implemented successfully in the system. Also, various parameters like efficiency, power factor and Total Harmonic Distortion (THD) have been compared for all the three converters. The simulation of the model is performed on the MATLAB/ SIMULINK software. By the given comparison of the obtained result, the apt converter can be chosen by the designer according to the specifications.**

Keywords—Total Harmonic Distortion (THD), Full Bridge Diode Rectifier (DBR), Brushless Direct Current Motor (BLDC), Power Factor Correction (PFC)

I. INTRODUCTION

Currently, the usage of BLDC motors has been increasing rapidly due to its various attractive characteristics like its increased dependability, wide range of speed, trouble-free maintenance and lesser noise while operation. These are generally used with DBR that allows it to withdraw non-sinusoidal current containing high input THD and low power factor. To overcome these difficulties a PFC converter is used in the design. The conventional BLDC drive is shown in Fig. 1 and the THD obtained is around 54% for that drive.

Various PFC DC-DC Converter topologies are available but the three topologies namely SEPIC, Cuk and Zeta converters are selected because of their wide range of voltage control, high efficiency and any polarity of output according to the chosen converter compared to the conventional Buck-Boost converter. Voltage follower approach is used for maintaining required constant output voltage from the PFC converter.

All the three converters can operate as buck or boost converters in accordance with their duty cycle. The Cuk converter like the Buck-Boost converter offers the negative output voltage polarity whereas both the SEPIC and Zeta converters offer the same voltage polarity as the input. Discontinuous Conduction Mode (DCM) is chosen for the operation to obtain less THD and unity power factor. Finally, the power factor, efficiency and THD of all the three converters are calculated and a comparison has been made.

Section 2 in the paper describes the system configuration taken for simulation. DC-DC converters are discussed in the section 3, 4 and 5. Section 6 discusses the three phase inverter used in the simulation and then voltage follower approach is described in Section 7. In section 8 modelling, simulation, and results of the controlled drive are discussed and in section 9 the conclusion is provided.

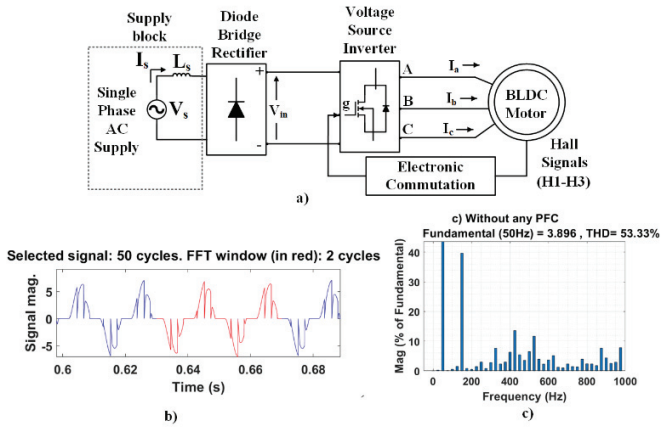


Fig 1. a) Block Diagram of the model without any PFC converter b) Input Current waveform "I_s" c) THD in the I/P current waveform without any PFC

II. SYSTEM CONFIGURATION

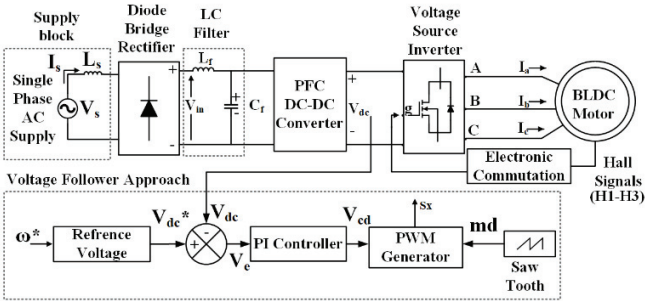


Fig 2. Block diagram of the SIMULINK model

In Fig. 2a PFC converter controlled by the voltage follower approach is feeding a BLDC motor. All the three PFC

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:01:01 UTC from IEEE Xplore. Restrictions apply.

converters contains a high frequency MOSFET, a diode 2 inductors and a coupling capacitor. The commutation process of the employed inverter is done electronically using the three Hall sensors so that the MOSFETs of the Voltage Source Inverter (VSI) operate at its fundamental frequency and thus reducing the switching losses. All the PFC converters in the paper operate in the DCM mode of operation such that the input inductor current in each converter is discontinuous and the voltage follower approach is used to control the DC link output voltage of the converter. The resulting waveforms of all the PFC DC-DC converters are obtained in the DCM mode and a small source inductance (L_s) is also taken at the supply side to improve the THD in the input current waveform and a LC filter is taken after the DBR to reduce the ripples in the obtained DC voltage.

III. CUK CONVERTER

The schematic for the CUK converter is shown in figure 3. The input inductor (L_x) acts as a filter for the DC input supply and hence prevents any large harmonic content to pass through.

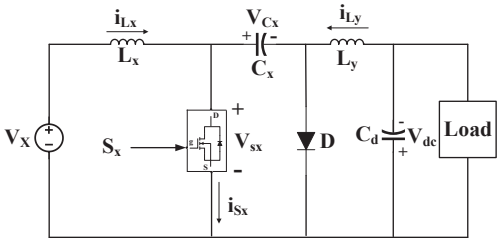


Fig 3. CUK converter schematic

A. Working priniciple of the CUK converter

During the ON period (Fig. 4a) of the high frequency MOSFET the DC voltage source (V_x) charges the input inductor (L_x) while the coupling capacitor (C_x) provides power to the load.

In the OFF period (Fig. 4b) of the MOSFET the input inductor (L_x) will discharge its stored energy through the diode (D) into the coupling capacitor and this will also charge the DC link capacitor (C_d) and provide power to the given load.

By performing volt-sec balance on the equations obtained in the respective ON and OFF states we get equation (1)

$$\frac{V_{dc}}{V_x} = \frac{D}{1-D} \tag{1}$$

The inductance and capacitance values of the converter are evaluated using the geometry in the current and voltage waveforms.

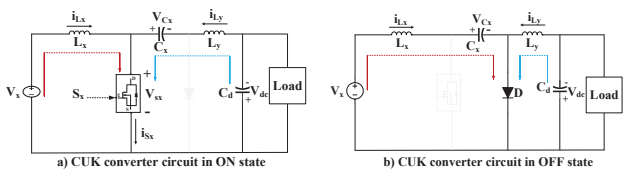


Fig 4. CUK converter circuits in a) ON state b) OFF state

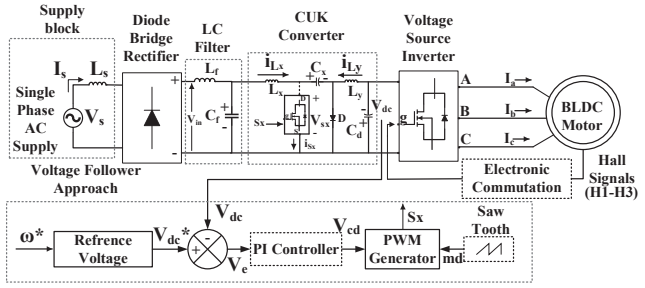


Fig 5. Block diagram of the SIMULINK model employing CUK converter

IV. SEPIC CONVERTER

In Fig.6 the schematic for the SEPIC converter is shown. This converter has the same number of energy elements and diode as that in the CUK converter, but all those elements are arranged in a different manner such that the SEPIC converter is also able to provide DC voltage greater or smaller than the supply voltage but with the same polarity as the source unlike the CUK converter.

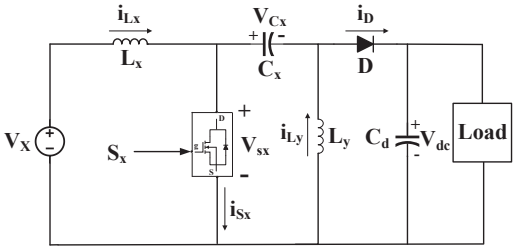


Fig 6. SEPIC converter schematic

A. Working priniciple of the SEPIC converter

In the ON period (Fig. 7a) of the switch (S_x) input inductor (L_x) is charged from the supply (V_x) and the output inductor (L_y) is charged from the coupling capacitor (C_x). The load during this time is supplied power from C_d and the output voltage can be approximately treated as constant during this period.

In the OFF period (Fig. 7b) of the S_x the load is fed through the supply V_x charging the capacitor C_x and discharging the inductor L_x in the process. The capacitor C_d is charged during this period from the discharge of L_y.

By writing voltage equations of the inductor in the respective ON and OFF states and then performing volt-sec balance we can obtain equation (2) which is same as the CUK converter but with same polarity as that of input.

$$\frac{V_{dc}}{V_x} = \frac{D}{1-D} \tag{2}$$

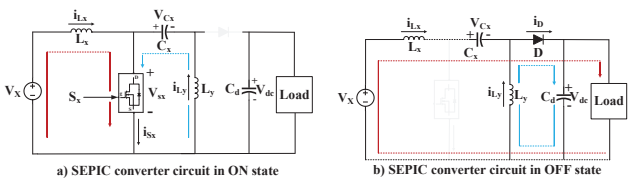


Fig 7. SEPIC converter circuits in a) ON state b) OFF state

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:01:01 UTC from IEEE Xplore. Restrictions apply.

TABLE I. DESIGN PARAMTERS FOR THE PFC CONVERTERS

	CUK Converter	SEPIC Converter	ZETA Converter
Input inductance "L_x"	$D*V_{in} / (\Delta I_{Lx})\ f_s$	$D*V_{in} / (\Delta I_{Lx})\ f_s$	$D*V_{in} / (\Delta I_{Lx})\ f_s$
Output inductance "L_y"	$V_{dc}(1-D) / (\Delta I_{Ly})\ f_s$	$V_{dc}(1-D) / (\Delta I_{Ly})\ f_s$	$V_{dc}(1-D) / (\Delta I_{Ly})\ f_s$
Coupling Capacitance "C_x"	$D / \{(Rf_s)\ (\Delta V_{Cx} / V_{dc})\}$	$D / \{(Rf_s)\ (\Delta V_{Cx} / V_{dc})\}$	$D / \{(Rf_s)\ (\Delta V_{Cx} / V_{dc})\}$
DC Link Capacitance "C_y"	$(1-D) / \{(8L_yf_s^2)\ (\Delta V_{Cy} / V_{dc})\}$	$D / \{(Rf_s)\ (\Delta V_{Cy} / V_{dc})\}$	$(1-D) / \{(8L_yf_s^2)\ (\Delta V_{Cy} / V_{dc})\}$

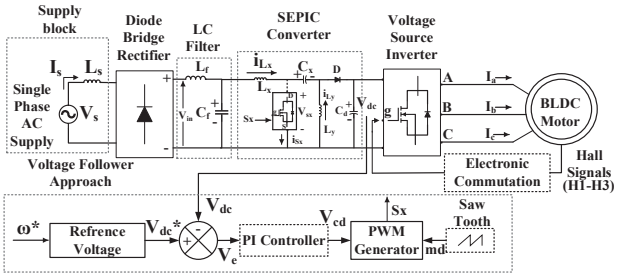


Fig 8. Block diagram of the SIMULINK model employing SEPIC converter

V. ZETA CONVERTER

In Fig. 9 the schematic for the ZETA converter is shown and like the previous two converters this also has two inductors (L_x, L_y) and two capacitors (C_x, C_y) arranged in a different fashion as shown in figure 9.

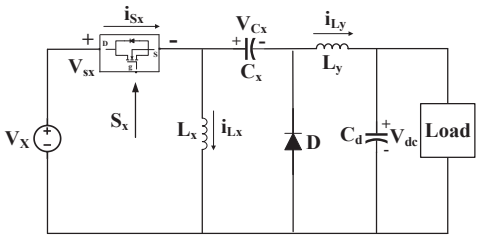


Fig 9. ZETA converter schematic

A. Working priniciple of the ZETA converter

In the switch ON period (Fig. 10a) the inductors L_x and L_y gets charged through the supply V_x and they store energy during this interval.

In the OFF period (Fig. 10b) of the S_x the inductors L_x and L_y discharge their stored magnetic energies into the DC link capacitor (C_d) and the load. The output voltage V_{dc} is almost constant with a small ripple because of the switching of the MOSFET.

By performing the volt-sec balance on the voltages across the inductor in the respective ON and OFF states we get the input-output relationship (eqn 3) such that the polarity of the O/P voltage is same as that of the input.

$$\frac{V_{dc}}{V_x} = \frac{D}{1-D} \qquad (3)$$

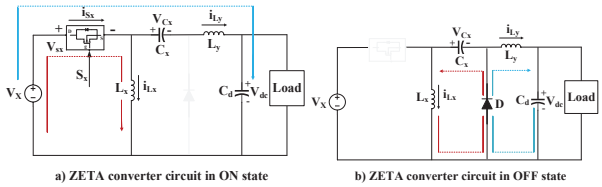


Fig 10. ZETA converter circuits in a) ON state b) OFF state

VI. THREE PHASE VOLTAGE SOURCE INVERTER

The 3 Φ inverter (figure 12) converts the DC output voltage of the PFC converter into smooth AC waveform at a corresponding frequency, it has six switches which complement each other i.e. in a particular leg only one switch can be ON at a time.

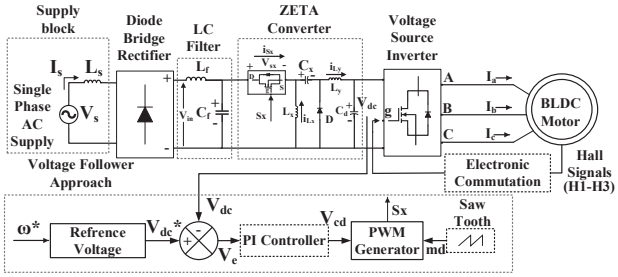


Fig 11. Block diagram of the SIMULINK model employing ZETA converter

The switching of the inverter is done by sensing the position of the BLDC motor through the three hall sensors that are present in the BLDC motor.

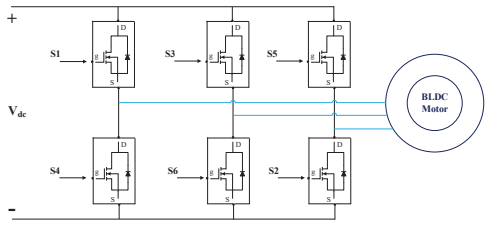


Fig 12. Three phase voltage source inverter

VII. VOLTAGE FOLLOWER APPROACH

All the three PFC converters are operating in the discontinuous mode i.e. the current in the "L_x" is discontinuous and so the best control scheme to control the voltage is the voltage follower approach.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:01:01 UTC from IEEE Xplore. Restrictions apply.

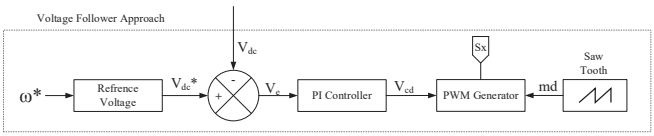


Fig 13. Block diagram representation of the voltage follower approach

A reference speed (ω* = 209.33 rad/sec) is taken and corresponding to this ω* a reference voltage (V_{dc}^* = 200V) is generated using

$$V_{dc}^* = kN^* \quad (4)$$

V_{dc}* is then compared with the instantaneous output of the PFC converter and corresponding to that a voltage error is generated, this error (V_e) is then given as input to the tuned PI controller and then the output (V_{cd}) is compared with a saw tooth waveform of high frequency such that a PWM signal for the switch of the PFC converter is generated.

VIII. Simulation and Results

The above models were simulated in the SIMULINK software and the following waveforms (Fig 14, 15, 16 & 17) were obtained after employing the respective PFC converters.

A. BLDC and PFC converter waveforms

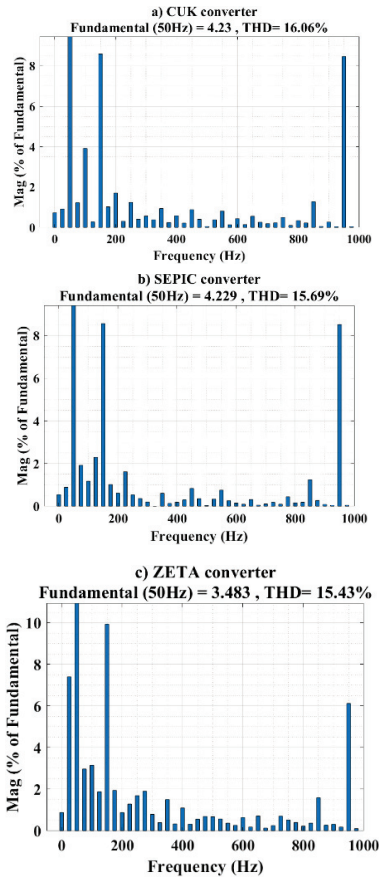
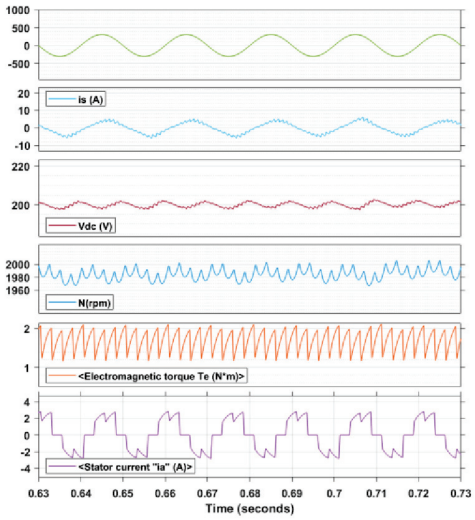
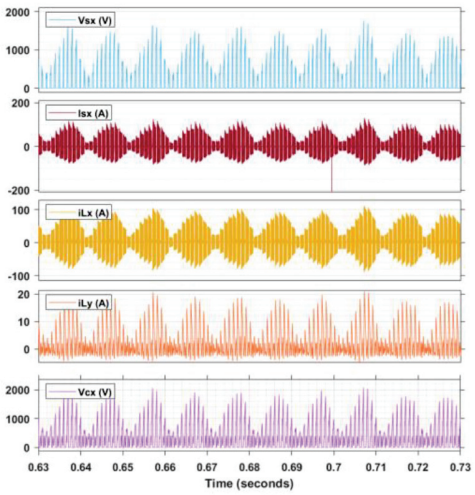


Fig 14. THD observed in the input current waveform a) CUK converter b) SEPIC converter c) ZETA converter

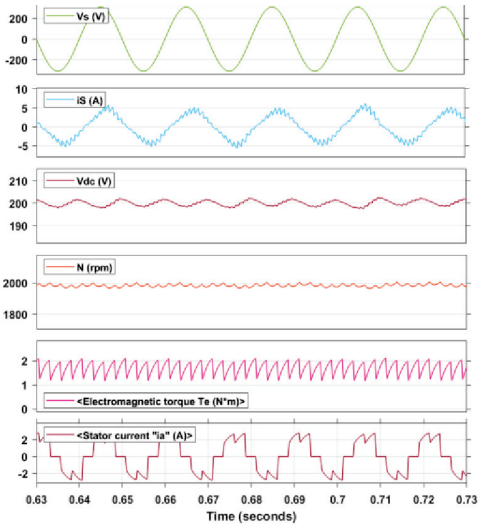


a)



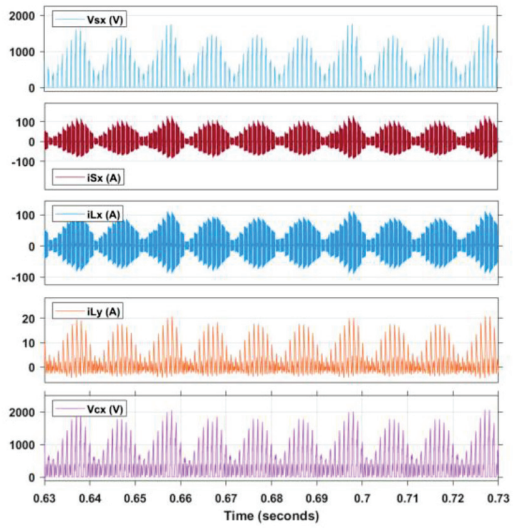
b)

Fig 15. a) BLDC waveforms b) CUK converter waveforms



a)

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:01:01 UTC from IEEE Xplore. Restrictions apply.



b)

Fig 16. a) BLDC waveforms b) SEPIC converter waveforms

B. *Efficiency calculations and comparison*

TABLE II. INDIVIDUAL EFFICIENCIES IN THE CUK CONVERTER MODEL

Device	Input Power (W)	Output Power (W)	Power efficiency
1. Full wave DBR	594.6	541.87	91.13%
2. CUK converter	541.87	427.8	78.94%
3. Inverter	427.8	422	98.64%

TABLE III. INDIVIDUAL EFFICIENCIES IN THE SEPIC CONVERTER MODEL

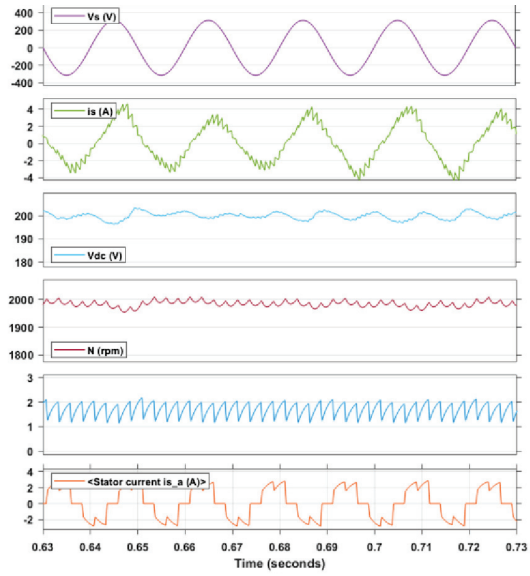
Device	Input Power (W)	Output Power (W)	Power efficiency
1. Full wave DBR	593.6	529.245	89.15%
2. SEPIC converter	529.245	427.21	80.72%
3. Inverter	427.21	422.1	98.80%

TABLE IV. INDIVIDUAL EFFICIENCIES IN THE ZETA CONVERTER MODEL

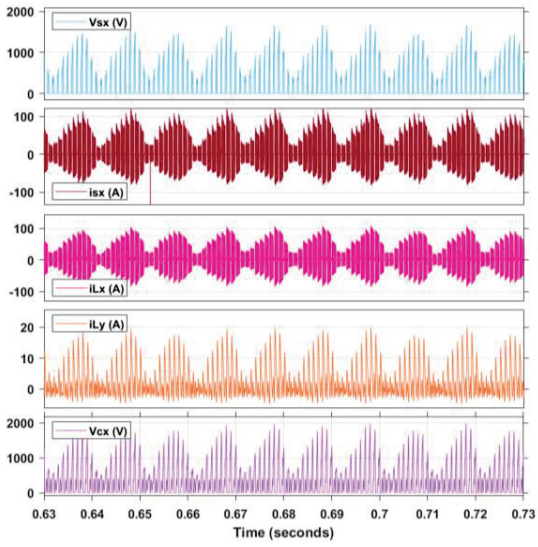
Device	Input Power (W)	Output Power (W)	Power efficiency
1. Full wave DBR	439.4	430.3	97.92%
2. ZETA converter	430.3	426	99%
3. Inverter	427.21	422.1	98.80%

TABLE V. PARAMETERS COMPARSION

Parameter	No Converter	CUK	SEPIC	ZETA
1. THD	53.4%	16.06%	15.69%	15.43%
2. Power Factor	0.7	0.93	0.91	0.93
3. Efficiency	88.19%	70.95%	70.34%	94.98%



a)



b)

Fig 17. a) BLDC waveforms b) ZETA converter waveforms

IX. CONCLUSION

The three non – isolated PFC DC-DC converter topologies viz Cuk, SEPIC and Zeta for the improvement of power quality in PMBLDC motors have been implemented and their in depth - analysis has been made successfully. The validation of the design of these topologies have been successfully done using the SIMULINK software. These topologies can be used for the applications that require low power around 500W. The selection of one of the topologies among these depends on the various parameters like THD, power factor, efficiency etc. These results on various topologies will provide the needed reference for the users and manufacturers of PMBLDC drives.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:01:01 UTC from IEEE Xplore. Restrictions apply.

ACKNOWLEDGMENT

We hereby acknowledge CONIT Publisher, annexed sponsors and Delhi Technological University for providing a good platform and consideration to showcase our findings. We would also like to express our sincere gratitude and appreciation for making the research work accomplished by providing enlightening opportunity and support even during unprecedented crisis of COVID-19 pandemic.

REFERENCES

[1] B. Singh, B.P Singh., S. Dwivedi.: 'Improved power quality AC–DC flyback converter for variable speed permanent magnet synchronous motor drive'. Proc. IEEE Power India Conf., New Delhi, India, April 2006, p. 8

[2] P. Venma, P. Vandana, "PV fed BLDC Motor using Zeta Converter with Minimized Torque Ripple for Water Pumping", International Conference on Emerging Technological Trends(ICETT), 2016.

[3] B. Singh, S. Singh, "Single-phase power factor controller topologies for permanent magnet brushless DC motor drives," IET Power Electron., 2010, Vol. 3, Iss. 2, pp. 147–175

[4] N. Siddharthan, B. Balasubramanian, "Performance Analysis of SEPIC, Luo and Zeta Converter", International Journal of Power Electronics and Drive Systems, March 2019.

[5] S. N. Pavithra, Dr. S. Umamaheshwari," Zeta converter fed BLDC motor for power factor correction and speed control", Conference on Emerging Devices and Smart Systems (ICEDSS), March 2016.

[6] V. Bist, B. Singh: 'PFC Cuk Converter- Fed BLDC Motor Drive', IEEE Transactions on Power Electronics, Volume: 30, Issue: 2, Feb. 2015; Page(s): 871 – 887

[7] Daniel W. Hart, 'Power electronics', Tata McGraw-Hill Education, 2011

[8] G. K. Dubey, 'Power Semiconductor Controlled Drives', Prentoce Hall, New Jersey, 1989

[9] B. Chandan, P. Dwivedi, S. Bose: 'An Experimental Study of SEPIC Converter with BLDC Motor As Application',IEEE-HYDCON, 2020

[10] S. Singh, G. Bhuvaneshwari, B. Singh: 'A Power Factor Corrected Two Stage Switched Mode Power Supply', IEEE International Conference on Power Electronics, Drives and Energy Systems, 16-19 Dec. 2012

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:01:01 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.

Potent Real-Time Recommendations Using Multi-Model Contextual Reinforcement Learning

Anubha Kabra, Anu Agarwal, and Anil Singh Parihar, *Member, IEEE*

***Abstract*— Widespread digitalization has led to almost all utilities and services thrive on an online medium. A real-time, personalized, and trend grasping recommendation system is necessary to enhance user experience and boost business on E-commerce platforms. We propose the Multi-Model Contextual Reinforcement Learning (MMCR) constituting three novel features for real-time and customized recommendations. The first feature is user-item interactive state embedding which uses not only item information but also assigns weightage to this information according to its usage history. It gives higher importance to the newly clicked items by the users than the older ones. Second, we devised Contextual Cluster Exploration (CCE) strategy. This strategy enhances the item-choice recommendations by consistently reducing the randomness during exploration. The third novelty is an item-based multi-agent framework that can tackle the case of sparsely chosen items. Generally, such items are disregarded in a single agent model as the more popular items take supremacy. Our technique ensures that the user-item history per item is learned separately; thus, no item is neglected. MMCR has shown an average of 5% increase in CTR rate. Moreover, CCE exploration gives a considerably higher score than state-of-the-art exploration strategies. Thorough experimentation demonstrates that our proposed strategy has shown significantly improved results over various state-of-the-art strategies.**

***Index Terms*— Deep Q-learning, human–computer interaction, multi-agent reinforcement, recommender systems, reinforcement learning (RL).**

NOMENCLATURE

Q	Item model.
R	Reward function.
a	Action.
s	State.
r	Reward.
U	User set.
E	Experience buffer.
i_h	User specific item history.
U_i	User features.
C_i	Context features.
l	True label.
N	Number of candidate items.
thresh	Threshold (Default : 0).
k	Number of items to be recommended.
LR	Logistic regression.
W&D	Wide and deep.
FM	Factorization machines.
DRN	Deep reinforcement learning framework for News recommendation.
DRR	Deep reinforcement learning based recommendation.
UCB	Upper confidence bound.
TS	Thompson sampling.
B	Bootstrapped.
AGT	Adaptive greedy threshold.
AE	Active exploration.
AAG	Active adaptive greedy.

Manuscript received October 12, 2020; revised May 9, 2021 and July 16, 2021; accepted July 21, 2021. *(Anubha Kabra, Anu Agarwal, and Anil Singh Parihar contributed equally to this work.) (Corresponding author: Anil Singh Parihar.)*

Anubha Kabra was with the Machine Learning Research Laboratory, Department of Computer Science and Engineering, Delhi Technological University, New Delhi 110042, India. She is now with Adobe Systems, Noida 201304, Uttar Pradesh, India (e-mail: anubhakabradtu@gmail.com).

Anu Agarwal was with the Machine Learning Research Laboratory, Department of Computer Science and Engineering, Delhi Technological University, New Delhi 110042, India. She is now with the Samsung Semiconductors India Research and Development Centre, Bengaluru 560037, Karnataka, India (e-mail: anuag7699@gmail.com).

Anil Singh Parihar is with the Machine Learning Research Laboratory, Department of Computer Science and Engineering, Delhi Technological University, New Delhi 110042, India (e-mail: parihar.anil@gmail.com).

Digital Object Identifier 10.1109/TCSS.2021.3100291

I. INTRODUCTION

IN TODAY'S fast-paced commerce, efficient recommendations to user bases have become a priority for most industries. From media visuals to advertising and E-commerce, the recommendation of items holds a central place. Hence, to balance with a continuously changing environment, user base, and market inclinations, a real-time, competent, and personalized system is necessary. For instance, Amazon, Facebook, Google, and most other big and small firms earn a large amount of their revenue from recommending advertisements [1]. Customer retention by showing the customer the right recommendation at the right time becomesthe ultimate goal of the competitive marketplace. Over time, a plenitude of ideas have been proposed to solve the real-time item recommendation problem. A variety of static techniques such as modifications of collaborative filtering (CF) [2]–[4], certain hybrid methodologies [5] have been proposed. However, these are static and do not perform efficiently in real-time scenarios. Static recommendation strategies have a fixed training and testing set that restricts it from capturing user behavior changes over time. The recommendations are shown to the users, while testing is solely dependent on the information learned during training. For instance, if a user prefers type A items during training, however, changes their preferences to type B later, the model might still be recommended type A items as it did not learn the changes. It ignores the stochastic nature of the user. Fig. 1 reflects the dynamic nature of two real time

2329-924X © 2021 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 01,2021 at 07:03:21 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.

datasets which we have used for experimentation. We plotted occurence of five random items against time for each dataset. The occurrence of an item strictly differs with timeline. Hence, training on a fixed segment of data cannot guarantee that the same trend will follow during testing. In Fig. 1, for TalkingData dataset, if the training covers the items 1, 3 and 4, and the trend shifts toward items 2 and 5 post training, the items recommended to users during testing would be 1, 3 and 4 instead of 2 and 5.

Content-based recommendation is a separate class of recommendation algorithms which leverage the user and item features [6], [7]. A segment of these algorithms are developed on top deep learning networks [8], [9], since they can co-relate non-linear user-item interactions more accurately. However, many content-based techniques [6], [9] suffer from the cold start problem. This problem occurs due to insufficient information available to make initial recommendations (when new users enter the system). Thus, random recommendations are shown to the user. It is crucial to solve the cold start problem due to the continuously changing user base in a real-time environment. Looking at another topic of research, reinforcement learning (RL) has been gaining popularity in gaming [10], [11], automation, and robotics [12], [13]. The general norm in an RL-based problem is to have an agent/agents in an environment, according to which it takes actions. The goal is to take actions in the environment which maximizes the long-term reward. The task of generating recommendations can be viewed as an RL problem, the purpose being to maximize customer retention and provide a quality user experience [14]–[16].

Another major challenge observed in the recommender system literature is utter neglect of the user-bases who do not always prefer mainstream or popular content [15]. Preferences of such users can be captured according to item-specific user space segregation. RL techniques have a module to explore or exploit. Effective exploration can advance the working of an RL model. State-of-the-art exploration techniques may result in high randomness [17], which adversely affects the model. Sometimes, exploration strategies take a long time to learn a trend [18]. In a fast-paced, real-time scenario, this negatively affects the recommendations.

Hence, the current state-of-the-art techniques have one or more of the following challenges.

1) Fails at dynamic front and hence fails to give real-time recommendations.
2) Suffers from cold-start problem.
3) Neglects unconventional item choices.
4) Fails to perform optimal yet effective exploration.

To tackle all the aforementioned challenges in delivering dynamic and personalized recommendations, we propose Multi-Model Contextual Reinforcement Learning (MMCR) Farmework. In order to address the issue of model stationarity, MMCR uses RL-based Deep Q-Networks. Its advantages are twofold. First, the use of deep neural networks enables efficient representation of all underlying non-linearities present between dataset features. Second, RL leverages on the sequential nature of the problem and grasps any observable change in the user preferences via continuous updates and feedbacks.

To address the problem of new users and users with unconventional choices, MMCR has a provision for separate models corresponding to every item in the itemset. It enables each model to identify the characteristics of its audience properly. Thus, whenever a new or user with unconventional choice is encountered, multi-model framework is able to provide relevant recommendations based on user-characteristics. It also ensures that popular items do not have an unfair edge over the unpopular ones. Furthermore, we propose a novel Contextual Cluster Strategy for Exploration. It works on the principle that we only explore random actions in an unforeseen situation. Clusters of user-context embeddings are formed to make the implementation of this principle feasible and efficient. As our knowledge of the environment gradually increases, the number of random items in the recommendation list decreases.

In this article, any subsequent mention of the term environment refers to the complete set of users and items available in the dataset. The system refers to our MMCR framework and the term model corresponds to the per-item neural network. State refers to the embedding generated using user features, session context features, and user's item history. The terms arm and action used interchangeably in the text indicate items in the itemset. When a user requests the system for recommendations, his/her specific features along with session-context features serve as the input. This input is further processed by the state generator (SG) module to obtain state embedding. Based on this state, the system selects k most relevant items and presents them to the user as output to obtain feedback. However, only the clicked item obtains actual feedback from the user in the form of a positive reward. All obtained feedbacks are stored in the system memory to be replayed on the respective models after fixed intervals.

We summarize our novelty and key contributions as follows.

1) Proposed a novel multi model framework which maintains an independent agent per item. Such a framework gives this architecture an edge over traditional recommendation techniques by integrating users with unconventional item choices.
2) Introduced a novel Contextual Cluster Exploration (CCE) for exploration which provides an adaptive exploration rate by forming clusters based on existing knowledge of the environment. CCE outperformed considerably on traditional exploration strategies.
3) Proposed a first of its kind method of incorporating user's interaction history with the system, into the input state. This enables the model to learn from user's previous interactions effectively.
4) Leveraged the benefits of both user and item features to recommend relevant items.
5) Demonstrated the effectiveness of our framework by performing thorough experimentation on two publicly available real-world datasets.

The rest of this article is structured as follows: Section II defines the Background and Related work, Section III gives the Problem Definition, Section IV focuses on Proposed work and our Pipeline for MMCR technique. Section V has com-

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:03:21 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.

prehensive Experimentation and Results evaluation. Section VI concludes this article and discusses the scope of future work in this domain.

II. Background and Related Work

This section is further divided into three subsections based on the recommendation strategy used. The first subsection summarizes the prior art, which treats recommendation as a supervised learning problem. Second lists all prominent strategies based on single-agent RL, and the last one describes the literature that has treated it as a multi-agent RL problem.

A. Traditional Recommendation Strategies

A significant amount of research has been done to solve the problem of generating recommendations that maximize the cumulative gain. The initial literature for recommendations describes the systems that are purely based on item popularity [19]. A critical class of recommenders that followed was content-based filtering [6], [7], which takes into account the correlations between the user and item features. Another major breakthrough was achieved with the advent of CF [3], [4], [20]. It can be further categorized as user-based, item-based, and hybrid filtering. In user-based filtering, recommendations are made based on the popular choices of the most-similar users [21]. Item-based filtering [22] shows the items which are highly similar to the ones that the user has previously selected. Other variant approaches such as Matrix Factorization [23], Factorization Machines [24], [25], Logistic regressions models have also managed to give a performance boost to the CF-based methods. The incorporation of Deep Neural networks in Recommenders is used for efficiently modeling the complex non-linearities [26]. Wang *et al* [27] uses denoising auto-encoders on top of Matrix Factorization model to extract useful features from the side information available. The key feature that distinguishes our model from the rest of these techniques is our framework's ability to learn at every step of the training as well as the testing. Hence, it can learn all the dynamic trends from the incoming data, unlike the traditional technique whose learning is static and is only limited to training data patterns.

B. Single Agent RL-Based Recommendation

In a generic single-agent RL setting, the agent observes a state s_t from the environment at timestep t. It uses a policy π to obtain the best action for the given state s_t. Action is pushed to the environment, and a reward r_t is obtained, leading to the next state s_{t+1}. The objective here is to tune the parameters θ of the policy π to obtain the maximum cumulative expected reward. RL frameworks can be broadly divided into value-based methods and policy-based methods. Both the methods involve the system in a current state s_t. γ refers to the discount factor.

1) Value-Based RL: Value-based RL involves the estimation of an action-value function. This function determines the expected sum of discounted future reward for a specific action a_t corresponding to the current state s_t. The Q-value function (Q) described above can be expressed mathematically using the Bellman equation [28] as follows:

$$Q(s) = E\big[r_{t+1} + \gamma\, Q(s_{t+1}) | (s_t = s)\big]. \tag{1}$$

Osband *et al.* [29] proposed DQN which suggested the use of deep neural networks for estimating the Q-function. The loss function (L) which is minimized by DQN is as follows:

$$L(s_t, a_t) = \left(r_t + \gamma \max_{a'} Q\big(s_{t+1}, a', \theta'\big) - Q(s_t, a_t, \theta)\right)^2. \tag{2}$$

Here, θ refers to the parameters of the target model and θ' refers to the parameters of the online network. The target model is used to generate the Q value predictions for a certain fixed number of iterations. After these iterations, target model's weights are updated to θ'.

Several attempts have been made to leverage Q-Learning [30], DQN frameworks for effective recommendations [31], [32]. Zheng *et al.* [15] uses DQN as a base model with Dueling Bandit Gradient Descent Strategy for exploration. User-Profiling has also been used as a value-based RL problem for generating recommendations [33]. Zhao *et al.* [34] leverages long-term and session-based knowledge using adversarial training and RL for recommending movies.

2) Policy-Based RL: Policy-based RL is especially helpful in the case of continuous state and action representations which cannot be efficiently dealt with using Q-Learning or DQN. It uses gradient ascent to optimize the parameters for maximizing the expected reward. In policy-based methods such as actor-critic [35], the actor accepts the state as input and outputs continuous action value. The action is further passed to the critic network along with the input state vector to generate Q value for the state-action pair. The gradient J for actor-critic is described as follows:

$$\Delta J(\theta) = E_\pi\left[\sum_{t=0}^{T} \Delta_\theta \log \pi_\theta(a_t | s_t; \theta) R\right] \tag{3}$$

where R is the sum of gamma-discounted rewards. Furthermore, the system generates a list of recommendations consisting of all top-scoring items. Several recommendation strategies [14], [36]–[38] use actor-critic as a base RL model with modifications in state representation, exploration functions, etc. to provide better results.

C. Multi-Agent RL-Based Recommendation

Multi-Agent RL strategy consists of multiple RL agents trying to maximize the discounted cumulative long-term reward collectively. However, with multiple agents in action simultaneously [39]–[41], each agent only has a partial view of the environment. Most common approach followed in multi-agent systems is decentralized execution and standard Q-value function evaluation [42]. However, this involves a significant issue of credit assignment for each agent. The global action performed is usually a combined result of execution performed by all agents together. Hence, the global action receives a global reward that cannot be linked to any one agent in particular [43]. Huang *et al.* [44] compares the global reward

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:03:21 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.

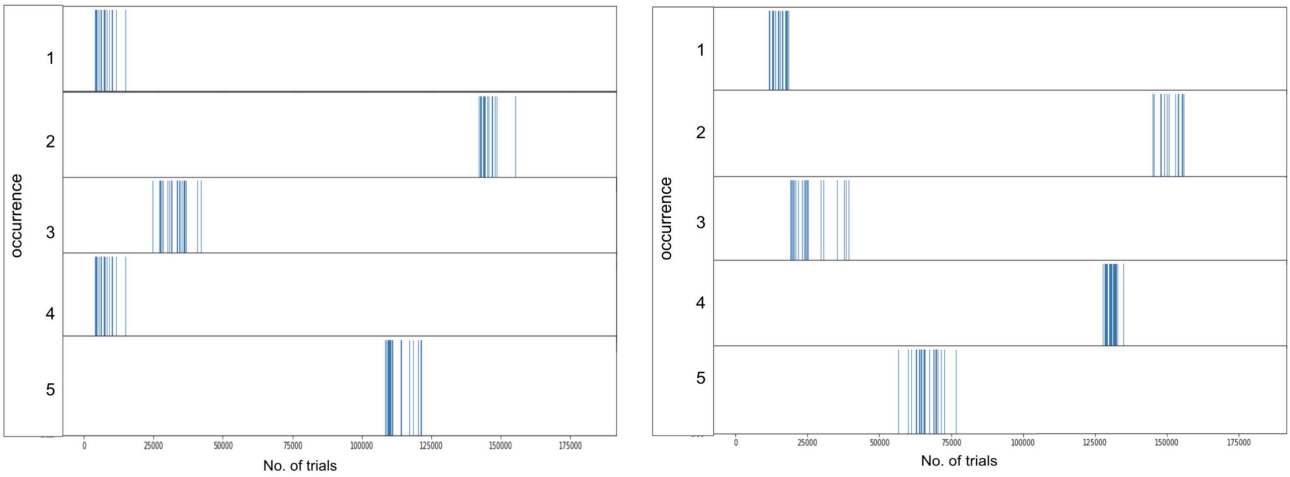


Fig. 1. Occurrences of random five items throughout the timestamp for TalkingData and Globo data, respectively.

with the reward obtained by taking a default outcome for all other agents except the one for which the reward is being calculated. Independent Q-learning [45] models multi-agent by learning wholly separate and independent DQN networks with the environment as the only source of interaction between them.

However, one problem that remains is the convergence and stability of the system. Since all other agents are updating their weights simultaneously, no single agent can consider its environment stationary, and hence it ends up chasing an ever-moving target. Most of the available literature on multi-agent systems deal only with the game-playing scenario and not with recommendation systems. Our approach involves using a separate agent corresponding to each available action (which is possible in a recommendation system scenario as there are only a discrete number of items available to choose from) in an attempt to overcome the issues mentioned above.

III. Problem Definition

Whenever a user i with user features u_i and context features c_i makes a recommendation request to the system, our model generates a list of top-k most relevant recommendations for the user. These recommendations are based on the results obtained from N RL agents, where N is the fixed number of items in the available itemset. The notations used throughout this article are summarized in the Nomenclature section.

IV. Proposed Work

In this section, we elaborate extensively on segments of our proposed architecture.

The novelty of the proposed work is as follows.

1) Developing a novel multi-model architecture for recommendation systems using per-item agent strategy with independent credit assignment.
2) SG Module is the first of its kind technique to incorporate the user interaction history and user features as the input. This enables the model to learn from user's previous interactions effectively.
3) CCE is a new method to perform enriched exploration by relying on its present knowledge of the environment to make smart and less randomized choices.

We first explicate the SG Module, Novel exploration strategy (CCE), and Multi-Agent RL to define concepts used in our framework. We then elucidate on our Model Architecture which incorporates all the aforementioned segments.

A. SG Module

This module leverages the extra information present in the commonly available user-interaction logs such as user, item, and context features. Fig. 2 shows its overall working. Its functioning can be broadly divided into two parts.

1) *Maintaining User's Interaction History:* Simple aggregation of all the item-labels that the user has previously clicked not only consumes excessive storage but also provides very little insight into user behavior. Instead, we propose to maintain interaction-history in terms of item-feature vectors. Hence, whenever a new item a is to be added in a user's history vector h_u, it is updated as follows:
$$h_u = \alpha h_u + a. \tag{4}$$
Here, α is the age multiplier used to decrease the impact of the old feature values as they move further into the past and to prevent feature values from exploding.
2) *Generating State Embeddings:* Auto-encoder is used at this step for generating uniform embeddings. The user and context features from the input request are concatenated with the interaction history feature vector i_h.
$v = u_t + c_t + i_h$. This concatenated vector v is fed as an input to the autoencoder. The output from the autoencoder is the required state embedding s_t.

B. CCE: Contextual-Cluster Exploration Strategy

For deciding between exploration and exploitation, we propose a novel strategy called CCE. This is diagrammatically

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:03:21 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.

Algorithm 1: Context Clustering Exploration

```
Initialize:
c_thresh
Random n- experience clusters
Define the centroid of each cluster randomly (c)
for each user state embedding s_t do
    d_min = MAX_VAL
    c_sel = None
    for 1 to n Cluster Centroids do
        d = Euclidean_dis(s_t, c_i)
        if d < d_min then
            d_min = d
    end
    if d_min > d_thesh then
        EXPLORE
        Recommend k random items to user
    else
        EXPLOIT
        Recommend k best items to user
    end
     After showing the user k items
    for j in k items do
        if Reward(j) = 1 then
            s'_t = Updated State Embedding
        else
            s'_t = Original State Embedding
        end
    end
    for 1 to n Cluster Centroids do
        d = Euclidean_dis(s'_t, c_i)
        if d < d_min then
            d_min = d
            c_sel = c_i
    end
    if d_min > d_thesh then
        Assign a new cluster to this state embedding
        c_count + = 1
        if c_count > c_thresh then
            Reassign all the clusters
    else
        Add state embedding (s'_t) to c_sel
    end
end
```

represented in Fig. 3. As described in the Algorithm 1, we first initialize centroids of n-Experience Clusters randomly. The default value of n is 10. For every input request at time t, the distance d_i of the generated state embedding s_t is calculated from all the current centroids using Euclidean distance. If the minimum distance ($d_{\min}$) is greater than the decided threshold (d_{thresh}), then we explore else exploit. During exploration, random k items are shown to user else best k items are chosen. After showing the items, the state embedding for the current user is updated according to the items clicked by him/her (s'_t). The distance of s'_t is calculated from all the existing clusters. If s'_t has its minimum distance ($d_{\min}$) smaller than (d_{thresh}), then s_t gets assigned to the nearest cluster. Otherwise, a new cluster is formed. However, if the total number of clusters exceed c_{thresh} (default: 15), then re-clustering of all the existing user state embeddings is done. A clear representation of the same is shown in Fig. 1(b).

C. Multi-Agent Reinforcement

Instead of letting a single agent handle the execution of all the incoming recommendation requests, we have divided the execution task among multiple homogeneous DQN-based agents. Unlike other multi-agent strategies, our framework differentiates the functioning of these agents in terms of the information they act upon. Each agent is responsible for learning the user-behavior or characteristics corresponding to a specific item/action. This not only makes these models independent of each other but also addresses one of the main issues faced by other multi-agent recommendation strategies, i.e., instability [46]. Other strategies are prone to instability as one agent cannot assume the environment to be stationary if other agents are getting updated simultaneously. In our case, since every agent acts on action-segregated information, they tend not to get affected by the update of weights in other models. Another major problem in current systems is the local credit assignment for global reward [43]. When multiple agents cumulatively score the candidate items and finally present a list, it is difficult to assign credit for the obtained reward. However, since our framework assigns a separate agent for every action, the credit assignment is entirely straightforward. Although multi-agent systems are highly prevalent in applications such as game-playing, their usage in recommender systems is still rare.

A significant advantage of our MMCR framework over other common single-agent recommendation strategies [14], [15] is that it gives relevant recommendations, even for the sparsely active users. Each model pays attention to only specific information and naturally does a better job at identifying what user and context combination suits the given item the best.

Moreover, MMCR does not neglect relatively unpopular items. As observed in several real-life scenarios, some items are highly popular among the users in general while others are liked by a rather specific set of users. In traditional recommender systems [5], [9], [15] however, the unpopular items might never get recommended to the right user. MMCR has the ability to learn the preferred user characteristics for a particular item, as it maintains a separate model for each item. In this framework, we have opted for an item-centric approach rather than the prevalent user-centric model. Thus, even if an item appears less frequently than others, its chances of being recommended to the right user are not hurt.

D. Model Framework

Our Multi-Model Contextual Reinforcement (MMCR) framework combines user, item, and user-item context features along with the user's positive interaction history to create a personalized and interactive recommender system. Instead of a single agent, the task of generating recommendations is spread over multiple RL agents. A detailed depiction of the proposed framework is given in Fig. 4. The step-by-step procedure for generating recommendations, following Algorithm 2 is as follows.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:03:21 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.

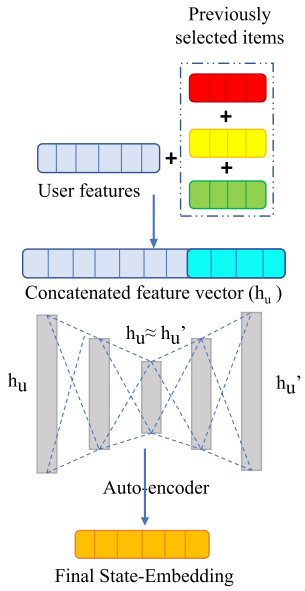


Fig. 2. SG module.

1) At each timestep t_i, SG (see Section IV-A) is fed with user and context features as input. This module combines them with the user's positive interaction history and feeds it to an auto-encoder, thus generating a uniformly sized state embedding s_t.
2) The state embedding s_t from the previous step is then sent to the Explore module (see Section IV-B) which decides whether to use exploration or exploitation for the current request. In this module, s_t's proximity is calculated from the existing Experience Clusters to decide if a similar state has been encountered previously or not. If the module decides that a similar state has previously been handled, the system proceeds with exploitation, and the state is assigned to one of the experience clusters. Otherwise, it explores the existing itemset randomly. This step produces a list L of k recommended items.
3) The list L generated from the previous step is pushed to the user to obtain feedback. Since we know the actual feedback of only one item in the itemset (i.e., the item clicked by the user), the agent corresponding to that item is updated with positive feedback for the current state embedding. The reward is defined as

$$R(s_t, a_t) = \begin{cases} 1, & \text{if } a_t = \text{right_arm} \\ 0, & \text{if } a_t \neq \text{right_arm}. \end{cases} \quad (5)$$

Here, s_t refers to current state, a_t refers to the selected item and right_arm is the correct item with respect to the state (s_t). Agents corresponding to the rest of the items on L are updated with negative feedback. Each embedding-item-reward tuple (s_t, a_t, r) is added to the Experience buffer memory. The buffer collects the embedding-item-reward tuples.
4) Once experiences equivalent to batch size have accumulated in the experience buffer, selective samples are picked up as per the algorithm and replayed on their respective agents. (In our case, we have given zero weightage to the negative experiences however it can be varied as per use-case by changing the value of thresh), as shown in Algorithm 2. This is performed in order to ensure that the model is more inclined toward the trends observed recently as compared to those observed in the distant past. Moreover, we update the Experience clusters by moving the cluster centroids according to the newly accumulated states using the k-means clustering algorithm. The Experience buffer is then emptied.

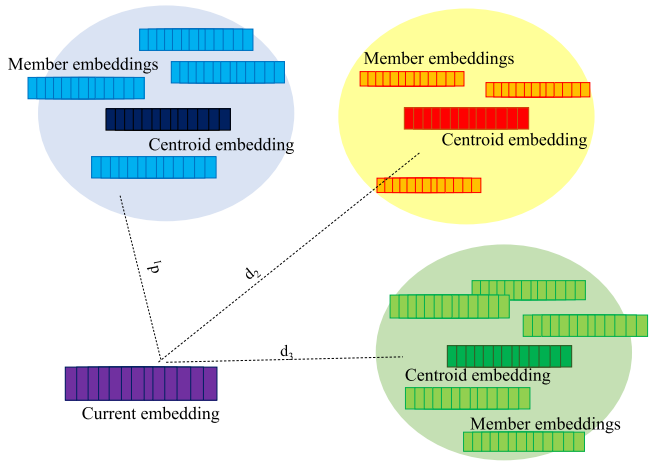


Fig. 3. Context-clustering exploration.

V. Experimentation and Results

A. Datasets

1) TalkingData User Gemographics Dataset: This dataset[1] contains user and mobile application interaction logs. The Data is collected from TalkingData SDK integrated within mobile apps. Here, the mobile application consumer is the proposed user, and mobile applications refer to the proposed itemset.

2) Globo.Com News Dataset: This dataset[2] has page view logs corresponding to news articles from a popular news portal of Brazil. (Globo.com). Here, the newsreaders are proposed users and news articles refer to the proposed itemset.

Keeping in mind several non-repetitive or hardly repetitive items as well as the vastness of the data, we chose to select most frequently appearing 40 items. The selection was made similarly for both the datasets to maintain the consistency of experiments. The data had a combination of categorical and continuous features which were one hot encoded and normalized, respectively. The statistics of both the datasets after pre-processing are given in Table I.

B. Baselines

1) Non-RL Baselines:

1) *LR:* Logistic Regression is widely used in the industry as a common baseline method due to its easy implementation and high efficiency. It takes all the contextual

[1]https://www.kaggle.com/c/talkingdata-mobile-user-demographics/data
[2]https://www.kaggle.com/gspmoreira/news-portal-user-interactions-by-globocom

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:03:21 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
DETAILS ABOUT THE DATASETS USED. HERE # REPRESENTS THE COUNT

Dataset	#Size	#Features	#Continuous	#Categorical	#User &Context Features	#Item Features
TalkingData	72,405	13	8	5	11	2
Globo	1,82,494	10	6	4	8	2

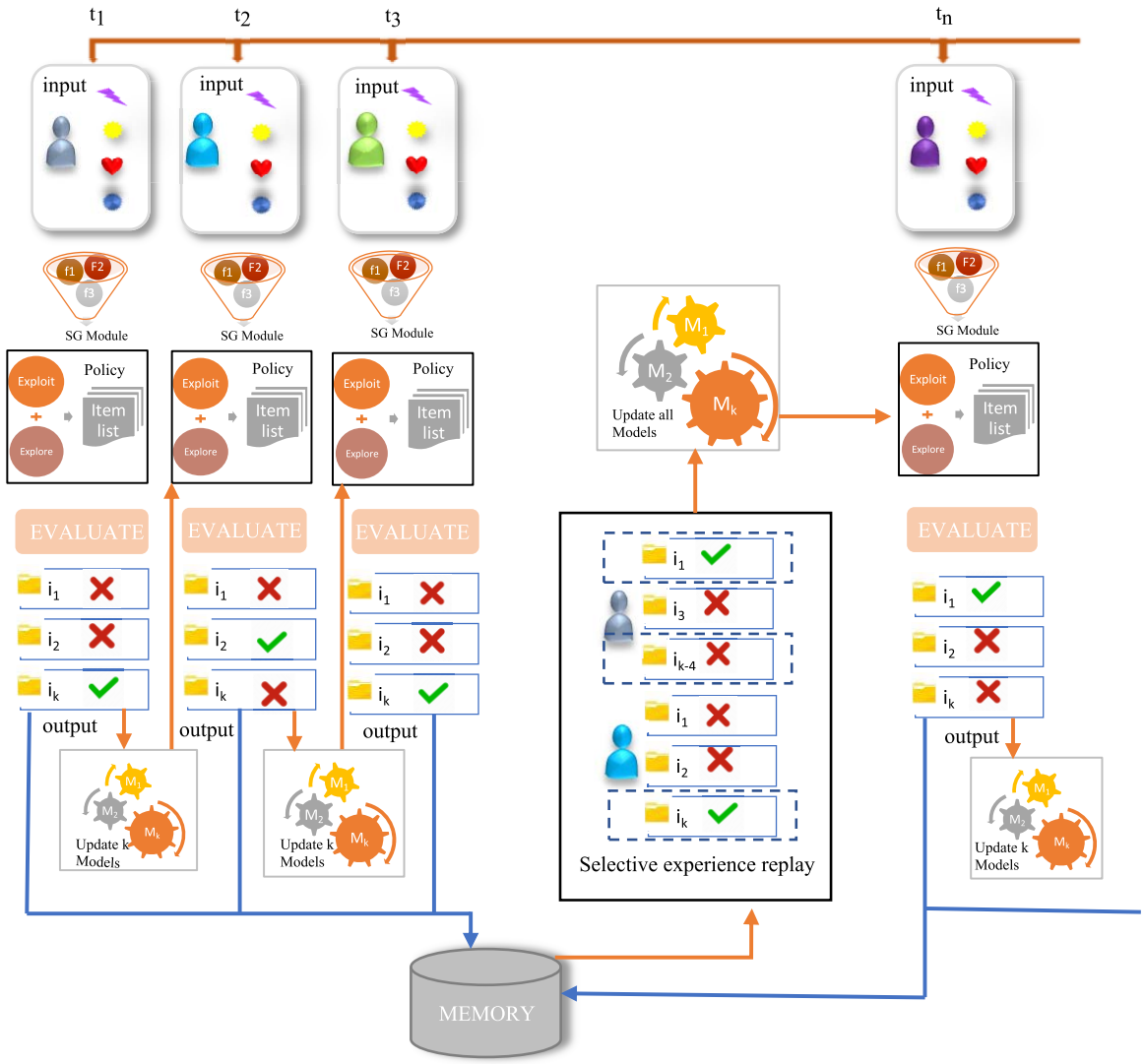


Fig. 4. MMCR pipeline.

features of users as input. This approach, however fails to model the complex non-linearities of the data features.

2) *W&D:* Wide & Deep [9] is a widely used state-of-the-art deep learning model combining the memorization (through a logistic regression on wide combinations of categorical features) and generalization (through a deep neural network embedding of the raw features) to predict the click label.
3) *FM [24], [25]:* Factorization Machine is a combination of factorization models and support vector machines (SVMs). This is advantageous for parameter estimation in sparse settings.[3]

[3]https://www.csie.ntu.edu.tw/ b97053/paper/Rendle2010FM.png

2) Single Agent RL:

1) *DRN [15]:* Deep Reinforcement Learning Framework for News Recommendation is a strategy that takes both user and item features into consideration. It uses double deep-Q networks DDQNs) as a base framework along with Dueling Bandit Gradient Descent-based exploration strategy.
2) *DRR [14]:* This stands for Deep RL-based Recommendation. A sequential regulation framework that has Actor-Critic RL-based pipeline is used for mapping relations between the users and recommendation system. Moreover, explicit user item interactions are used while modeling.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:03:21 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.

Algorithm 2: Overall Methodology

```
Initialize:
input : User Features (u), Context Features(c),User
 specific item History(i_h), Item Features (a) , E= [null] ,
for i E (1 to N) initialize DQN Models Q_i
for session i = 1 : data.length do
    arms = []
    right_arm = l_i
    U_i = (u_0, u_1, ....., u_m)
    C_i = (c_0, c_1, ....., c_m)
    i_h = [null]
    s_i = State_Generator(U_i, C_i, i_h)
    if (Explore Module(s_i) == True) then
        EXPLORATION
        for k in K do
            arms.append(random(arm_list))
        end
    else
        EXPLOITATION
        for j in Items do
            scores.add(Q_j.predict(s_i))
            Sort scores in descending order
            arms.add(top k items in scores)
        end
    end
    Update i_h with right_arm
    s'_i = State_Generator(U_i, C_i, i_h)
    for m in arms do
        if (m == l_i) then
            reward = 1
            Q_i.fit([s_i, m, reward, s'_i])
            E.add([s_i, m, reward, s'_i])
        else
            reward = 0
            if Random(num) < thresh then
                E.add([s_i, m, reward, s'_i])
            Q_i.fit([s_i, m, reward, s'_i])
        end
    end
    if (E.length % Batch_size == 0) then
        for (t in E) do
            Q_i.fit(t)
        end
        E = []
end
```

3) Exploration Strategies: Explorations constitute a major part of any Multi Arm Bandit-based recommender system. We combined our proposed model with various existing exploration strategies to compare their performance against CCE.

1) *Epsilon-Greedy [47], [48]:* This is the most basic strategy which explores arms at random if a randomly generated number has a value less than a chosen parameter ϵ, and prefers exploitation otherwise.
2) *UCB [48], [49]:* This leverages the uncertainty in action-value pair. A confidence interval for each item is formed using a suitable distribution which determines the confidence about the actual action-value of an item which lies in this bound. Results have been compared with Linear and Bootstrapped versions of UCB.
3) *TS [47], [49]:* It essentially involves selecting an arm with a probability proportional to it being the best arm. Results have been compared with Linear and Bootstrapped versions of TS.
4) *Adaptive Greedy (AG) [49]:* This is a variant of classic ϵ-greedy. However, during the execution ϵ is varied in a controlled manner. A threshold value for probability scores is generated. This threshold value is used to decide between exploration and exploitation. However, the threshold value changes when the number of trials exceeds a chosen window size so as to adapt the model to give better predictions. Two variants of this strategy namely AAG and AGT have been compared here.
5) *AE [49]:* This strategy uses active-learning as its base concept. It tries to select an observation depending upon the amount of leverage it gets on the model if its label were known.

C. Evaluation

1) Simulated Online Environment: The evaluation for all the methodologies was done in an online simulation. We have adhered the following guidelines.

1) The data were processed in increasing order of timestamp. This maintained the time-dependent pattern of the data to adhere to how users would enter the system in a real-time scenario.
2) Although the traditional methods perform separate training, our framework adopts learning while testing. Immediate feedback and experience replay provides continuous learning. This is elaborated in Section IV-D.

The simulation takes each state (depicting a user entering the system) as entry into the system. The system predicts the item for each state, which is then compared with the actual label (right arm), The right_arm is not known to the system before the comparison similar to a real-world scenario. For example: If a user enters a news website, k news items will be recommended to her. She can decide to choose one or none. If any of the items are clicked, it counts for a positive reward, else negative.

2) Online Evaluation Metrics: The following metrics were used for the evaluation of all the techniques compared in this article.

1) *CTR:* Click Through Rate is an effective measure for assessing the effectiveness of recommender systems in understanding the user click behavior. It is calculated for each item as follows:

$$\text{CTR}_i = \frac{\text{Total number of clicks}}{\text{Total number of impressions}}. \tag{6}$$

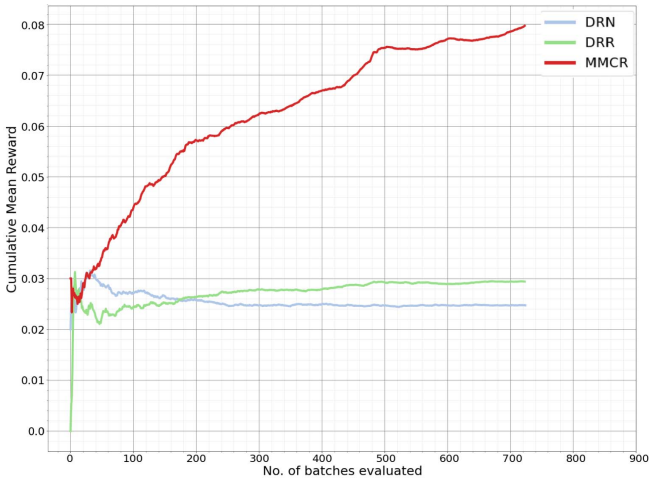
Further, average of CTR values of all the items is calculated to obtain the final CTR value as follows:

$$\text{CTR} = \frac{\sum_{i=0}^{N} \text{CTR}_i}{N}. \tag{7}$$

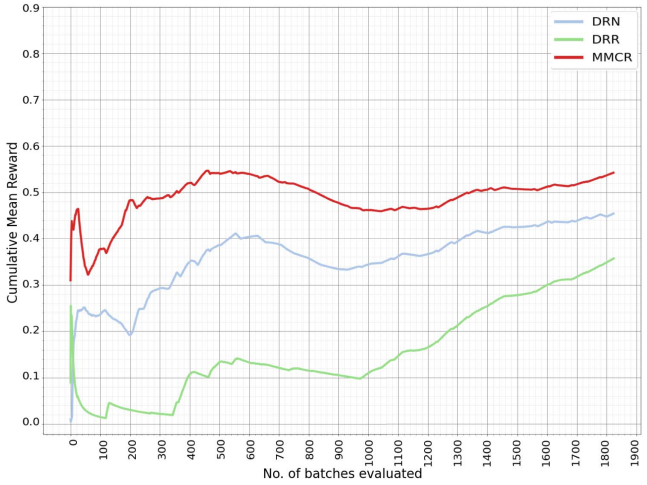
2) *Relative CTR:* It is a measure of evaluating how a recommendation strategy is performing on a given dataset

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:03:21 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.



(a)



(b)

Fig. 5. CMR trends on RL-based techniques for $k = 1$. The y-axis signifies CMR and x-axis signify the number of experiences. The red, blue and green line stands for MMCR, DRN and DRR, respectively. (a) TalkingData SDK. (b) Globo.

or platform as compared to other strategies when applied on the same dataset. The Relative CTR for ith strategy can be calculated as

$$\text{Relative CTR} = \frac{\text{CTR}_i}{\sum \text{CTR of other strategies}}. \quad (8)$$

3) *CMR:* Cumulative Mean Reward is a useful measure to continuously observe and analyze the performance of recommender systems over time. At any given point, if R is the number of rows evaluated so far, CMR is calculated as follows:

$$\text{CMR}_R = \frac{\sum_{i=0}^{R} \text{Reward}_i}{R}. \quad (9)$$

3) Offline Evaluation Metrics: Following popular recommender system evaluation measures have been compared for the offline evaluation of this article.

1) *Precision at k:* Precision is defined as the fraction of the total number of predicted recommendations (i.e., k) which are relevant to the given user

$$\text{Precision@k} = \frac{\text{No. of relevant recommendations}}{\text{Total no. of recommendations made}}. \quad (10)$$

2) *Recall at k:* Recall or sensitivity for k-recommended items is defined as the ratio of number of relevant recommendations to the total number of items relevant to the given user

$$\text{Recall@k} = \frac{\text{No. of relevant recommendations}}{\text{Total no. of clicked items}}. \quad (11)$$

The experimentation settings for the proposed method uses a deep neural network with two hidden layers for each item model. The batch size is 1000, i.e., the policy is updated after every trial, and experience replay after the batch size is reached. The optimizer used is Adam and loss is Binary Crossentropy.

D. Results

The comparison of the proposed techniques is made with the state-of-the-art recommendation strategies as well as the various exploration strategies mentioned in Section V-B. The exploration strategies are implemented on our base pipeline. Results are obtained in both online and offline settings. Tables IV and V show the performance of CCE as an exploration strategy over the rest. CCE majorly outperformed mentioned state-of-the-art techniques. Furthermore, the performance of MMCR in comparison to state-of-the-art RL and non RL-based strategies has been shown in Tables II and III. For maintaining consistency with the non-RL-based strategies, we have only shown results on the test set (i.e., 30% of the entire dataset). To have a comparison over the timestep for each RL-based technique, we plotted the CMR graphs shown in Figs. 5–7. The subsequent paragraphs highlight the intuition behind the observed trends.

1) *Context-Based Exploration:* MMCR focuses on user feature and context-based exploration. It can be observed from Tables IV and V that CCE performs best, irrespective of the number of items recommended (k) for both the datasets. CCE tends to shift gradually from exploration to exploitation as our knowledge of the environment increases. The positive effect of this shift can be seen in the significantly higher CTR, Precision, and Recall values for CCE as compared to other exploration strategies. Being a context-based strategy, Section IV-D deeply narrows the possibility of random items to be shown to the users. AAG and AE give comparable performance to CCE. These exploration strategies change the amount of exploration, generally reducing it as the model learns over time. However, AAG and AE do not use contextual information for this decay in exploration. The ϵ-greedy, UCB, and TS-based exploration do not do justice as ϵ-greedy relies on a consistent degree of randomness during exploration. UCB and TS require multiple instances of the same item to be shown before the exploration stabilizes. In a brief period, these two techniques can hurt performance.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:03:21 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 II
ONLINE AND OFFLINE EVALUATION FOR TALKINGDATA DATASET. BEST RESULTS ARE HIGHLIGHTED IN BOLD

Type	Method	Online Evaluation			Offline Evaluation					
		CTR			Precision			Recall		
		k=1	k=3	k=5	k=1	k=3	k=5	k=1	k=3	k=5
Non RL based	LR	0.0022	0.0042	0.0043	0.0045	0.0042	0.0036	0.0058	0.0050	0.0046
	W&D	0.0032	0.0010	0.0005	0.0051	0.0047	0.0042	0.0056	0.0052	0.0049
	FM	0.0036	0.0021	0.0014	0.0048	0.0039	0.0037	0.0059	0.0043	0.0041
Single Agent RL	DRN	0.210	**0.222**	0.192	0.462	0.413	0.407	**0.618**	0.573	0.583
	DRR	0.213	0.203	0.199	0.477	**0.459**	0.416	0.602	0.563	0.548
	MMCR	**0.231**	0.218	**0.206**	**0.518**	0.454	**0.439**	**0.618**	**0.603**	**0.594**

TABLE III
ONLINE AND OFFLINE EVALUATION OF GLOBO DATASET. BEST RESULTS ARE HIGHLIGHTED IN BOLD

Type	Method	Online Evaluation			Offline Evaluation					
		CTR			Precision			Recall		
		k=1	k=3	k=5	k=1	k=3	k=5	k=1	k=3	k=5
Non RL based	LR	0.0028	0.0031	0.0027	0.0053	0.0046	0.0040	0.0059	0.0055	0.0047
	W&D	0.0039	0.0027	0.0011	0.0058	0.0044	0.0043	0.0060	0.0056	0.0051
	FM	0.0036	0.0031	0.0019	0.0056	0.0052	0.0047	0.0066	0.0061	0.0053
Single Agent RL	DRN	0.321	0.219	0.200	0.399	0.387	0.381	0.542	0.527	**0.566**
	DRR	0.305	0.281	**0.211**	0.443	0.428	**0.408**	0.602	0.589	0.531
	MMCR	**0.374**	**0.288**	0.208	**0.509**	**0.431**	0.404	**0.611**	**0.603**	0.552

TABLE IV
VARIATION OF EXPLORATION STRATEGIES ON TALKINGDATA DATASET. BEST RESULTS ARE HIGHLIGHTED IN BOLD

Exploration Technique	Online Evaluation			Offline Evaluation					
	CTR			Precision			Recall		
	k=1	k=3	k=5	k=1	k=3	k=5	k=1	k=3	k=5
E-Greedy	0.228	0.138	0.108	0.406	0.399	0.392	0.488	0.431	0.416
Lin UCB	0.156	0.122	0.103	0.237	0.221	0.206	0.318	0.305	0.283
TS	0.161	0.117	0.092	0.261	0.245	0.218	0.324	0.320	0.291
B- UCB	0.212	0.147	0.106	0.366	0.347	0.329	0.419	0.405	0.394
B-TS	0.219	0.131	0.101	0.372	0.321	0.319	0.427	0.400	0.400
AGT	0.221	0.183	0.110	0.418	0.409	0.395	0.461	0.433	0.417
AE	0.226	0.195	0.115	0.452	0.438	0.424	0.519	0.508	0.500
AAG	0.229	**0.218**	0.117	0.488	**0.461**	**0.442**	0.606	0.593	0.585
CCE	**0.231**	**0.218**	**0.206**	**0.518**	0.454	0.439	**0.618**	**0.603**	**0.594**

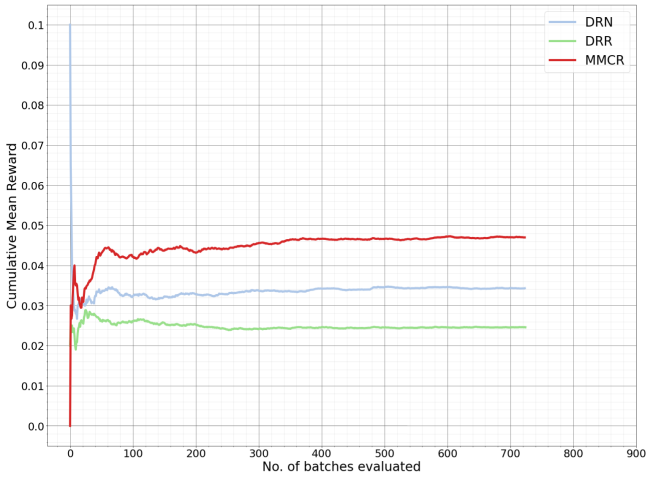
TABLE V
VARIATION OF EXPLORATION STRATEGIES ON GLOBO DATASET. BEST RESULTS ARE HIGHLIGHTED IN BOLD

Exploration Technique	Online Evaluation			Offline Evaluation					
	CTR			Precision			Recall		
	k=1	k=3	k=5	k=1	k=3	k=5	k=1	k=3	k=5
E-Greedy	0.286	0.251	0.222	0.471	0.432	0.421	0.589	0.531	0.520
Lin UCB	0.213	0.207	0.196	0.312	0.309	0.295	0.304	0.288	0.273
TS	0.224	0.209	0.200	0.328	0.314	0.306	0.319	0.308	0.297
B- UCB	0.236	0.218	0.203	0.331	0.329	0.308	0.345	0.321	0.318
B-TS	0.241	0.230	0.211	0.372	0.345	0.329	0.416	0.407	0.395
AGT	0.299	0.283	0.266	0.487	0.423	0.410	0.510	0.502	0.496
AE	0.329	0.315	0.307	0.488	0.420	0.401	0.572	0.543	0.521
AAG	0.362	0.310	0.309	**0.491**	0.428	0.394	0.609	**0.599**	**0.585**
CCE	**0.374**	**0.341**	**0.322**	0.477	**0.431**	**0.404**	**0.611**	**0.599**	0.552

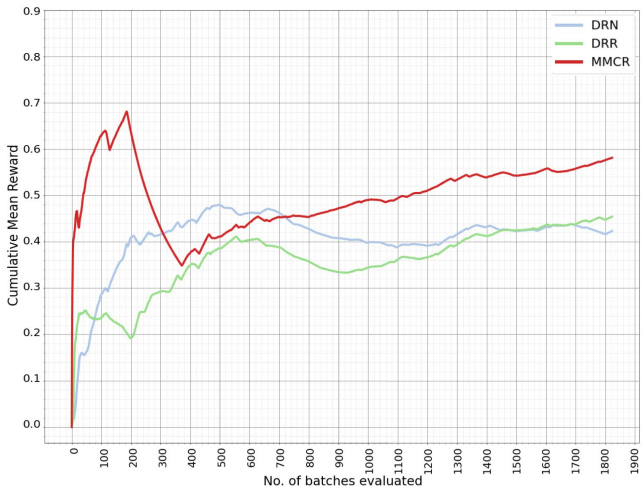
2) *Multi-Agent Versus Single-Agent:* When observing CMR in Figs. 5(a), 6(a) and (b), and 7(a) it is evident that DRN, DRR, and MMCR have a similar curve in the beginning. Over time, the performance of MMCR supersedes the other RL-based strategies. This trend can be accredited to the provision of item-based independent multi-agent pipeline of MMCR. Single-agent techniques eventually give less weightage to sparsely used items. This is a disadvantage in the case of recommendations as some items may have an acquired user set, e.g., a TV show recommender has generic shows as well as the ones which have a limited user base. In a single model-based technique such as DRR or DRN, the latter types of shows may get sidelined as they are not frequently chosen. With independent arm-based modeling, this problem is solved as the user type is easily recognized by the model of that particular item and thus item is correctly recommended. Overall, MMCR tends to

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:03:21 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.

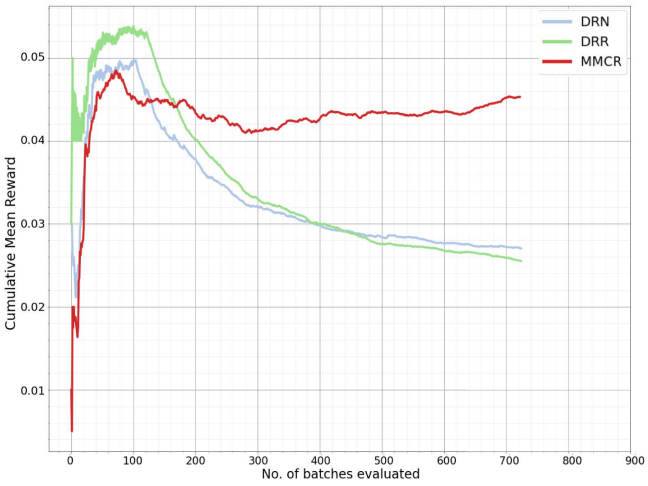


(a)

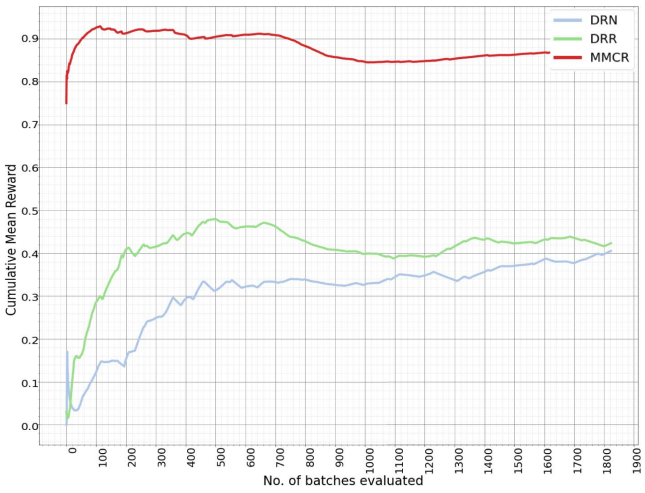


(b)

Fig. 6. CMR trends on RL-based techniques for $k = 3$. The y-axis signifies CMR and x-axis signify the number of experiences. The red, blue and green line stands for MMCR, DRN, and DRR, respectively. (a) TalkingData SDK. (b) Globo.



(a)



(b)

Fig. 7. CMR trends on RL-based techniques for $k = 5$. The y-axis signifies CMR and x-axis signify the number of experiences. The red, blue and green line stands for MMCR, DRN, and DRR, respectively. (a) TalkingData SDK. (b) Globo.

outperform DRN and DRR across both the datasets for all values of k (see Figs. 5–7). The same trend can also be reaffirmed from Tables II and III for online, as well as offline metrics.

3) *User-Item Weighted Interaction History:*The state embedding representation Section IV-A used by MMCR has used item information fruitfully by developing embeddings using the user-item interaction history as mentioned in Section IV-A. DRN and DRR use item information but do not give adequate weightage to the user-item interaction history. Item history for a particular user can give a deep insight into user behavior. MMCR ensures that new item features have a greater weight in interaction history than older items. Tables II and III as well as Figs. 5–7 representing CMR graphs have verified our claim by stating significantly improved results throughout for MMCR. In Figs. 5(a), 6(a), and 7(a) initially CMR graphs show that DRN and DRR are in competence with MMCR technique. In contrast for Figs. 5(b), 6(b) and 7(b) the MMCR technique has an edge from the beginning. Nevertheless, over time, MMCR captures the trends far better than DRR and DRN. All the techniques eventually achieve stability in CMR trends.

Although the current RL-based techniques such as DRN and DRR give improved results over non-RL-based strategies such as W&D and FM (see Tables II and III), they are conquered by MMCR. The non-RL techniques fail to perform in real-time data as the learning during training may not be sufficient to recognize the trends during testing. MMCR has three novel properties which are item-based multi-agent framework, non-conventional way of state embedding representation, and novel exploration CCE (see Section IV-D) which significantly improve its performance.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:03:21 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.

VI. Conclusion and Future Work

In this article, we propose an item-based multi-agent recommender system with the capability to tackle real-time dynamic recommendations with ease. This is especially beneficial to sparsely yet periodically occurring item sets that may be ignored in a single agent strategy. The basis of the technique relies on Contextual Multi-Arm Bandit with a novel exploration strategy called CCE which conquers SOTA exploration strategies by using effective clustering with minimal randomity. Moreover, using user and item information with an inclination to learn more from recent item trends using user-item interactive history has increased the potential for relevance in recommendations. The continuous updates and experience reply ensure current trends are learned and keep reinforcing them throughout. Our extensive experimentation has shown that these properties significantly improve results on both offline and online evaluation metrics. For future work, improving scalability by reducing the number of models in multi-agent RL or incorporating the properties into a single agent model has a scope. Moreover, new techniques of clustering can be tested for this CCE exploration strategy to make it advance.

References

[1] J-Clement. (2021). *Selected Online Companies Ranked by Total Digital Advertising Revenue From 2012 to 2020*. [Online]. Available: https://www.statista.com/statistics/205352/digital-advertising-revenue-of-leading-online-companies/

[2] A. S. Das, M. Datar, A. Garg, and S. Rajaram, "Google news personalization: Scalable online collaborative filtering," in *Proc. 16th Int. Conf. World Wide Web (WWW)*, 2007, pp. 271–280.

[3] B. Sarwar, G. Karypis, J. Konstan, and J. Reidl, "Item-based collaborative filtering recommendation algorithms," in *Proc. 10th Int. Conf. World Wide Web (WWW)*, 2001, pp. 285–295.

[4] J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen, "Collaborative filtering recommender systems," in *The Adaptive Web*. Berlin, Germany: Springer, 2007, pp. 291–324.

[5] J. Liu, P. Dolan, and E. R. Pedersen, "Personalized news recommendation based on click behavior," in *Proc. 15th Int. Conf. Intell. user Interfaces (IUI)*, 2010, pp. 31–40.

[6] M. Kompan and M. Bieliková, "Content-based news recommendation," in *Proc. Int. Conf. Electron. Commerce Web Technol.* Cham, Switzerland: Springer, 2010, pp. 61–72.

[7] M. J. Pazzani and D. Billsus, "Content-based recommendation systems," in *The Adaptive Web*. Berlin, Germany: Springer, 2007, pp. 325–341.

[8] L. Zheng, V. Noroozi, and P. S. Yu, "Joint deep modeling of users and items using reviews for recommendation," in *Proc. 10th ACM Int. Conf. Web Search Data Mining*, Feb. 2017, pp. 425–434.

[9] H.-T. Cheng *et al.*, "Wide & deep learning for recommender systems," in *Proc. 1st Workshop Deep Learn. Recommender Syst.*, 2016, pp. 7–10.

[10] D. Silver *et al.*, "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, 2016.

[11] G. Tesauro, "Temporal difference learning and TD-Gammon," *Commun. ACM*, vol. 38, no. 3, pp. 58–68, Mar. 1995.

[12] J. Kober, J. A. Bagnell, and J. Peters, "Reinforcement learning in robotics: A survey," *Int. J. Robot. Res.*, vol. 32, no. 11, pp. 1238–1274, 2013.

[13] S. Levine, C. Finn, T. Darrell, and P. Abbeel, "End-to-end training of deep visuomotor policies," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 1334–1373, 2015.

[14] F. Liu *et al.*, "Deep reinforcement learning based recommendation with explicit user-item interactions modeling," 2018, *arXiv:1810.12027*. [Online]. Available: http://arxiv.org/abs/1810.12027

[15] G. Zheng *et al.*, "DRN: A deep reinforcement learning framework for news recommendation," in *Proc. World Wide Web Conf. (WWW)*, 2018, pp. 167–176.

[16] A. Kabra, A. Agarwal, and A. S. Parihar, "Cluster-based deep contextual reinforcement learning for top-k recommendations," in *Proc. Int. Conf. Comput. Commun. Syst. (I CS)*, vol. 170. Shillong, India: Springer, 2021, p. 125.

[17] M. Tokic, "Adaptive ε-greedy exploration in reinforcement learning based on value differences," in *Proc. Annu. Conf. Artif. Intell.* Ho Chi Minh City, Vietnam: Springer, 2010, pp. 203–210.

[18] P. Auer, "Using confidence bounds for exploitation-exploration tradeoffs," *J. Mach. Learn. Res.*, vol. 3, pp. 397–422, Nov. 2003.

[19] H. Steck, "Item popularity and recommendation accuracy," in *Proc. 5th ACM Conf. Recommender Syst. (RecSys)*, 2011, pp. 125–132.

[20] Z. Cui *et al.*, "Personalized recommendation system based on collaborative filtering for IoT scenarios," *IEEE Trans. Services Comput.*, vol. 13, no. 4, pp. 685–695, Jul. 2020.

[21] F. Ge, "A user-based collaborative filtering recommendation algorithm based on folksonomy smoothing," in *Advances in Computer Science and Education Applications*. Berlin, Germany: Springer-Verlag, 2011, pp. 514–518.

[22] M. Deshpande and G. Karypis, "Item-based top-N recommendation algorithms," *ACM Trans. Inf. Syst.*, vol. 22, no. 1, pp. 143–177, 2004.

[23] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *IEEE Comput.*, vol. 42, no. 8, pp. 30–37, Aug. 2009.

[24] S. Rendle, "Factorization machines," in *Proc. IEEE Int. Conf. Data Mining*, Dec. 2010, pp. 995–1000.

[25] S. Rendle, Z. Gantner, C. Freudenthaler, and L. Schmidt-Thieme, "Fast context-aware recommendations with factorization machines," in *Proc. 34th Int. ACM SIGIR Conf. Res. Develop. Inf. (SIGIR)*, 2011, pp. 635–644.

[26] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Neural collaborative filtering," in *Proc. 26th Int. Conf. World Wide Web*, 2017, pp. 173–182.

[27] K. Wang, L. Xu, L. Huang, C.-D. Wang, and J.-H. Lai, "SDDRS: Stacked discriminative denoising auto-encoder based recommender system," *Cognit. Syst. Res.*, vol. 55, pp. 164–174, Jun. 2019.

[28] R. Bellman, *Dynamic Programming*, vol. 707. New York, NY, USA: Courier Corporation, 2013.

[29] I. Osband, C. Blundell, A. Pritzel, and B. Van Roy, "Deep exploration via bootstrapped DQN," in *Proc. Adv. neural Inf. Process. Syst.*, 2016, pp. 4026–4034.

[30] N. Taghipour and A. Kardan, "A hybrid web recommender system based on Q-learning," in *Proc. ACM Symp. Appl. Comput. (SAC)*, 2008, pp. 1164–1168.

[31] V. Mnih *et al.*, "Playing atari with deep reinforcement learning," 2013, *arXiv:1312.5602*. [Online]. Available: http://arxiv.org/abs/1312.5602

[32] X. Zhao, L. Zhang, Z. Ding, L. Xia, J. Tang, and D. Yin, "Recommendations with negative feedback via pairwise deep reinforcement learning," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2018, pp. 1040–1048.

[33] H. Liang, "DRprofiling: Deep reinforcement user profiling for recommendations in heterogenous information networks," *IEEE Trans. Knowl. Data Eng.*, early access, May 29, 2020, doi: 10.1109/TKDE.2020.2998695.

[34] W. Zhao *et al.*, "Leveraging long and short-term information in content-aware movie recommendation via adversarial training," *IEEE Trans. Cybern.*, vol. 50, no. 11, pp. 4680–4693, Nov. 2020.

[35] I. Grondman, L. Busoniu, G. A. D. Lopes, and R. Babuska, "A survey of actor-critic reinforcement learning: Standard and natural policy gradients," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 42, no. 6, pp. 1291–1307, Nov. 2012.

[36] L. Wang, W. Zhang, X. He, and H. Zha, "Supervised reinforcement learning with recurrent neural network for dynamic treatment recommendation," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2018, pp. 2447–2456.

[37] X. Zhao, L. Xia, L. Zhang, Z. Ding, D. Yin, and J. Tang, "Deep reinforcement learning for page-wise recommendations," in *Proc. 12th ACM Conf. Recommender Syst.*, Sep. 2018, pp. 95–103.

[38] X. Zhao, L. Zhang, L. Xia, Z. Ding, D. Yin, and J. Tang, "Deep reinforcement learning for list-wise recommendations," 2017, *arXiv:1801.00209*. [Online]. Available: http://arxiv.org/abs/1801.00209

[39] F. Buccafurri, D. Rosaci, G. M. L. Sarnè, and L. Palopoli, "Modeling cooperation in multi-agent communities," *Cognit. Syst. Res.*, vol. 5, no. 3, pp. 171–190, Sep. 2004.

[40] P. Tommasino, D. Caligiore, M. Mirolli, and G. Baldassarre, "A reinforcement learning architecture that transfers knowledge between skills when solving multiple tasks," *IEEE Trans. Cognit. Develop. Syst.*, vol. 11, no. 2, pp. 292–317, Jun. 2019.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:03:21 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.

[41] Y. Rizk, M. Awad, and E. W. Tunstel, "Decision making in multiagent systems: A survey," *IEEE Trans. Cognit. Develop. Syst.*, vol. 10, no. 3, pp. 514–529, Sep. 2018.
[42] J. Feng *et al.*, "Learning to collaborate: Multi-scenario ranking via multi-agent reinforcement learning," in *Proc. World Wide Web Conf. World Wide Web (WWW)*, 2018, pp. 1939–1948.
[43] J. N. Foerster, G. Farquhar, T. Afouras, N. Nardelli, and S. Whiteson, "Counterfactual multi-agent policy gradients," in *Proc. 32nd AAAI Conf. Artif. Intell.*, 2018, pp. 1–9.
[44] H. Huang, Q. Zhang, and X. Huang, "Mention recommendation for Twitter with end-to-end memory network," in *Proc. 26th Int. Joint Conf. Artif. Intell.*, Aug. 2017, pp. 1872–1878.
[45] A. Tampuu *et al.*, "Multiagent cooperation and competition with deep reinforcement learning," *PLoS ONE*, vol. 12, no. 4, Apr. 2017, Art. no. e0172395.
[46] L. Busoniu, R. Babuska, and B. De Schutter, "A comprehensive survey of multiagent reinforcement learning," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 38, no. 2, pp. 156–172, Mar. 2008.
[47] S. Agrawal and N. Goyal, "Thompson sampling for contextual bandits with linear payoffs," in *Proc. Int. Conf. Mach. Learn.*, 2013, pp. 127–135.
[48] J. Vermorel and M. Mohri, "Multi-armed bandit algorithms and empirical evaluation," in *Proc. Eur. Conf. Mach. Learn.* Cham, Switzerland: Springer, 2005, pp. 437–448.
[49] D. Cortes, "Adapting multi-armed bandits policies to contextual bandits scenarios," 2018, *arXiv:1811.04383*. [Online]. Available: http://arxiv.org/abs/1811.04383



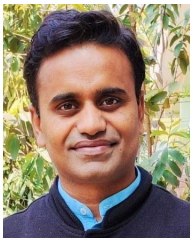
Anubha Kabra received the B.Tech. degree in computer science and engineering from Delhi Technological University, Delhi, India, in 2020.

She has worked with the Machine Learning Laboratory, Delhi Technological University, the Multimodal Digital Media Analysis (MIDAS) Laboratory, Indraprastha Institute of Information Technology (IIIT) Delhi, Delhi, and the Media and Data Science Research Laboratory, Adobe Systems, Noida, Uttar Pradesh, India, on various AI and data science-based projects. She joined Adobe Systems as a Software Development Engineer. Her current research interests include machine learning, deep learning, natural language processing, and reinforcement learning.



Anu Agarwal received the B.Tech. degree in computer science and engineering from Delhi Technological University, Delhi, India, in 2020.

She has worked with the Machine Learning Laboratory, Delhi Technological University. She is currently working as a Senior Engineer with the Advanced Sensor Algorithm Development Team, Samsung Semiconductors India Research and Development Centre, Bengaluru, Karnataka, India. Her current research interests include reinforcement learning, 3-D scene reconstruction, neural network optimization, computer vision, and deep learning.



Anil Singh Parihar (Member, IEEE) received the B.Tech. degree in electronics and communication engineering from Uttar Pradesh Techinal University, Lucknow, India, in 2005, the M.Eng. degree in electronics and communication engineering from Delhi College of Engineering, Delhi, India, in 2008, and the Ph.D. degree in computer vision and evolutionary computing from Delhi Technological University, Delhi, in 2016.

He joined the Department of Information Technology, Delhi Technological University, as an Assistant Professor, in 2010. He is currently an Associate Professor with the Department of Computer Science and Engineering, Delhi Technological University. His current research interests include computer vision, machine learning, deep learning, pattern recognition, soft computing, and evolutionary algorithms.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:03:21 UTC from IEEE Xplore. Restrictions apply.

Quantitative EEG Feature Selection by MajorityVoting for Alcohol Use Disorder Detection

Ruchi Holker
Department of Information Technology
Delhi Technological University
New Delhi, India
ruchi.holker@gmail.com

Seba Susan
Department of Information Technology
Delhi Technological University
New Delhi, India
seba_406@yahoo..in

2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI) | 978-1-6654-0358-0/21/$31.00 ©2021 IEEE | DOI: 10.1109/BHI50953.2021.9508483

Abstract— **Alcohol Use Disorder (AUD) is a social and health issue that is usually diagnosed by subjective tests. In this paper, we propose a method to identify significant Quantitative Electroencephalogram (QEEG) features that will aid in diagnosing AUD. We extract amplitude, spectral, range, and connectivity features, 7264 in all, from the four frequency bands associated with the 64 EEG channels. To reduce the complexity of the classification model, the top-3000 significant features are identified using a majority vote amongst different feature ranking methods: - ANOVA, chi-square test, Gini Index and Information gain ratio. SVM, KNN, Random Forest and ANN classifiers are used for the classification. Experiments on a benchmark dataset prove that our method performs better than the state of the art. Among all classifiers, the highest accuracy is 97.85% obtained by the ANN classifier for the proposed method.**

Keywords— ***Quantitative EEG features, Connectivity features, Amplitude features, Spectral features, Range EEG features, Feature selection, Majority voting, Alcohol Use Disorder***

I. INTRODUCTION

Electroencephalogram (EEG) is the most well-known technique for studying brain activity. EEG signals are neuropsychological signals, that can best capture real-time brain activity in a non-invasive manner, and present a good correlation between the complex brain structure and its activity. Alcohol Use Disorder (AUD) is a widespread social issue, that can cause damage to health and personal relationships in the long run. An alcoholic person can suffer various cognitive deficiencies such as short time learning and memory deficit problem [19], lack of decision making capacity, and behaviour changes that include depression.

In the past, major contributions to alcoholism identification and classification using EEG have been presented by various researchers. Greenblatt and Levin first proposed studying the correlation between EEG and chronic alcoholism using a mathematical model in 1944 [2]. After that, in 1980, Berry and Pentreath presented their paper on the neurophysiology of alcohol [3]. Neuroscience has since then been used to explain alcohol's acute and long-term effects on the nervous tissue. In the past few years, researchers have improvised on methodologies for classifying the alcoholic subject EEG and non-alcoholic subject EEG in time domain [10] [18]. A notable approach for classifying alcoholic and non-alcoholic data is to pick a subset of EEG electrodes for classification [14] [7]. An Improved Binary Gravitational Search Algorithm (IBGSA) is proposed as an optimization method for selecting the best EEG channels for rapid alcoholism screening in [9]. Several studies have explored empirical mode decomposition method to extract the features from the alcoholic and non-alcoholic EEG signal [5] [13]. Several features have been proposed over the recent years for the classification of alcoholic data. These vary from statistical features [8][12], and spectral features [20][15] to wavelet filter banks [4]. In [6], a set of Stationary features were proposed that capture frequency and amplitude properties, as well as features of inter-hemispheric connectivity. Connectivity features have also been proposed to analyze brain functioning [17]. The major contribution of this paper is the inclusion of range and connectivity features along with amplitude and spectral features for AUD detection. Significant features are then selected by a majority vote amongst state-of-the-art feature ranking methods to reduce the complexity of the classification model and improve the classification performance. The organization of this paper is as follows. The proposed method is presented in section II, the dataset is described in section III, the results are analyzed in section IV, and the final conclusions are given in section V.

II. PROPOSED METHOD

Quantitative EEG[11][16] involves application of mathematical programs to transform the raw EEG into frequency bands. First, a low pass filter at cut-off 30 Hz is used and then the EEG signal is down sampled at 128 Hz. The resultant EEG is divided into delta (0.5-4 Hz), theta (4-7 Hz), alpha (7- 13Hz), and beta (13-30Hz) frequency bands. Fig. 1 shows the framework of the proposed method.

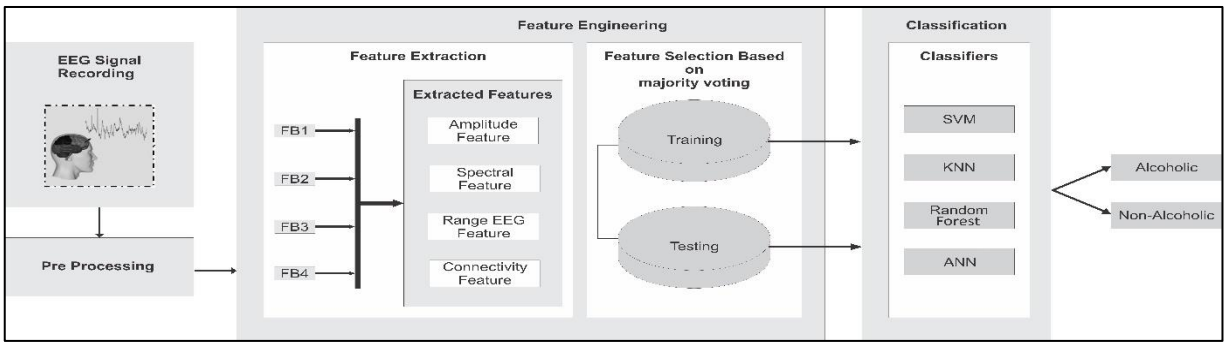


Fig. 1 Overview of the methodology

978-1-6654-0358-0/21/$31.00 ©2021 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:41:25 UTC from IEEE Xplore. Restrictions apply.

In this section, we describe the proposed methodology that finds the most prominent QEEG features for AUD detection, from the four frequency bands pertaining to the 64 EEG channels. We extract 1440 amplitude features, 1680 spectral features, 1920 rEEG features and 1200 connectivity features using bipolar montage. Another 1024 spectral features: AP, RP, Entropy and SD, we extract from CAR montage. A total of 7264 features are extracted in this manner from the four frequency bands pertaining to the 64 EEG channels. A majority voting scheme for an ensemble of feature ranking methods is used to short list the top-3000 features out of the total 7264 features. The selected features are classified by SVM, KNN, Random Forest (RF) and ANN classifiers. We first describe the different features used in the experimentation, followed by a description of the feature ranking ensemble used to identify the most significant QEEG features.

A. Feature Extraction

Twenty six different type of features Twenty six different types of features are extracted from the four frequency bands of each of the 64 electrodes; these are listed in Tables I to IV. Table I list the amplitude features. Amplitude is computed from signal power and signal envelope.

TABLE I: AMPLITUDE FEATURES

S.No	Measures	Description
1	Amp Total Power	Total power of a signal
2	Amp SD	Standard deviation
3	Amp Sk	skewnness
4	Amp Ku	Kurtosis
5	Amp env mean	Mean Value
6	Amp env SD	Standard deviation of envelope

Table II all the features derived from the spectral domain. EEG signal is converted into frequency domain using the Fast Fourier transformation, for frequency ranging from 0.5 to 30Hz at a resolution of 0.5 Hz. Then all the features are extracted from each frequency band for every channel.

TABLE II: SPECTRAL FEATURES

S.No	Measures	Description
1	AP	Absolute Power
2	RP	Relaetive Power
3	SP_En	Shannon Entropy
4	SP_Flatness	Wiener Entropy
5	SP_SD	Difference in spectral estimations for a short period of time
6	Edge_frq	Spectral edge frequency: 95% of spectral power contained between 0.5 and fc Hz (cut-off frq)
7	FD	Fractal dimension

Table III lists all the rEEG or range EEG (RE) [21] features (similar to amplitude-integrated EEG) used in the proposed method. rEEG is another way to measure the amplitude peak to peak voltage of a signal that is different from the previous amplitude measures [21] [16].

TABLE III: RANGE EEG FEATURES

S.No	Measures	Description
1	RE_mean	RE Mean
2	RE_medine	RE Median
3	RE_Low_mar	RE 5 percentile of lower margin
4	RE_Upper_mar	RE 95th percentile of upper margin
5	RE_Width	RE upper margin - lower margin
6	RE_SD	RE standard deviation
7	RE_CO	RE coefficient of variation
8	RE_asymmtry	RE measure of skew about median

Functional connectivity between two brain regions is de-fined by the statistical association between measures of activity recorded for the two regions in bipolar montage. Table IV lists the connectivity features used in our experiment.

TABLE IV: FUNCTIONAL CONNECTIVITY FEATURES

S.No	Measures	Description
1	BSI	Brain symmetry index
2	Conn_Correlation	Correlation between envelopes of hemisphere-paired channels
3	Conn_Coh_M	Coherence mean
4	Conn_Coh_Max	Coherence max value
5	Conn_Coh_frq	Coherence frequency of max value

B. Feature Selection

For finding the most important QEEG features that aid in AUD diagnosis, an ensemble of feature ranking methods isused. Mutual information based feature selection is a reliable technique but it is computationally expensive [22]. The following feature ranking techniques constitute our proposed ensemble:

- ANOVA: Analysis of variance (ANOVA) is a statistical model which determines the statistical difference between the independent groups.
- Chi-square test: The chi-square test find the relationship among categorical data. It is a statistical hypothesis test which estimates the features which are dependent on each other.
- Information Gain Ratio: Information gain ratio is the reduction in uncertainty. It is the ratio of information gain and attribute's intrinsic information.
- Gini index(GI): Gini index is a feature selection method which measures the inequality or discrimination level of the features with in a class.

C. Classifiers

In this paper, we have used four different classifiers to distinguish alcoholic and non-alcoholic EEG data. The four classifiers are Support Vector Machines (SVM), k-Nearest Neighbor (KNN), Random Forest (RF), and Artificial Neural Network (ANN). SVM classifies the data based on the margin maximization concept. KNN is a non-parametric classification algorithm that uses similarity measures to classify new cases among already stored available instances. RF is a non-linear ensemble classifier that contains N number of decision trees. Each decision tree has a different set of hyper-parameters on a sub-sample of a dataset. ANN is a complex adaptive classification system that can change its internal structure based on adjusting the weight of a connection. For the experiments and result analysis, we used MATLAB R2019a version.

III. DATASET

EEG data of alcoholic and non-alcoholic subjects were obtained from the University of California (UCI) [1]. The EEG data is extracted from 64 electrodes which are placed on the scalps of the subjects. The electrode position on subject scalps were located using the American

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:41:25 UTC from IEEE Xplore. Restrictions apply.

Electroencephalographic association standard. EEG signals of 120 subjects were recorded when they were shown 90 images of various objects. Each subject completed 120 trials for various stimuli. The signals have a resolution of 12 bits, and the EEG signals are measured at a frequency rate of 256 Hz for 32 seconds. After removing trails that had unwanted eyes and noise artifacts, the EEG recordings of each class were obtained. In this study, we used the "large dataset" version in the database, which contains separate training and testing data sets. The training and testing sets contain EEG signals of 10 alcoholic and 10 non-alcoholic subjects with 30 trials for each subject. Fig. 2(a) shows a single electrode signal of an alcoholic subject, and 2(b) shows a single electrode signal of a control subject.

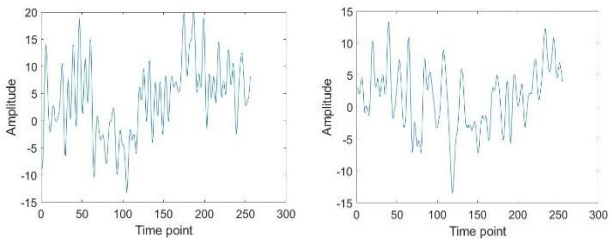


Fig. 2 (a) Alcoholic signal (b) Control signal

IV. RESULTS AND ANALYSIS

This section describes the main findings of our work. Extensive experiments are carried out to evaluate the proposed method for the classification of alcoholic and non-alcoholic EEG signals. Generally connectivity features are used by the neurophysiologists for the analytical studies [17]. In the proposed work, however, we include the connectivity features along with other domain features for classification. In our analysis, the feature vector contains 7264 features that were derived from four frequency bands pertaining to the 64 EEG channels and ranked individually by different feature selection and ranking methods such as ANOVA, information gain ratio, Chi-square and Gini index feature selection methods. Each feature selection method assigns a score to each feature according to their criteria. Majority voting is now performed among the four feature ranking methods. After that we select top 50, 500, 1000, 1500, 2000, 2500, 2867, 3000 and 3500 features selected by majority voting technique for classification. We find that before and after3000 features, there is not a noticeable change in accuracy scores. So top-3000 highly ranked features are selected for the classification. Table V summarizes the Accuracy, F1-score, Precision, and Recall results for the proposed method using all four classifiers (KNN, SVM, RF, and ANN). We compare our results with that of [15], [16], and [17] that are state-of-the-art methods for QEEG classification. The proposed method gives the maximum accuracy of 97.85% using ANN classifier. The other classifiers like SVM, KNN, RF achieve accuracy of 95.71%, 94.22%, and 91.91% respectively.

TABLE V: PERFORMANCE COMPARISON OF PROPOSED AND EXISTING METHODS FOR CLASSIFYING ALCOHOLIC AND CONTROL EEG SIGNALS

Author Name & Year	Model	Accuracy	F1-Score	Precision	Recall
Wajid et al [15], 2017	KNN	75.74%	0.7569	0.7604	0.7574
	SVM	84.32%	0.8432	0.8433	0.8432
	RF	79.37%	0.7936	0.7944	0.7937
	ANN	86.47%	0.8647	0.8647	0.8647
Toole et al [16], 2017	KNN	96.20%	0.9620	0.9629	0.9620
	SVM	96.53%	0.9653	0.9665	0.9653
	RF	93.40%	0.9338	0.9372	0.9340
	ANN	94.88%	0.9488	0.9502	0.9488
Huang et al [17], 2020	KNN	71.95%	0.7193	0.7198	0.7195
	SVM	80.69%	0.8055	0.8177	0.8069
	RF	77.56%	0.7750	0.7777	0.7756
	ANN	83.66%	0.8363	0.8385	0.8366
Proposed Method	KNN	94.22%	0.9428	0.9444	0.9422
	SVM	95.71%	0.9570	0.9590	0.9571
	RF	91.91%	0.9189	0.9235	0.9191
	ANN	97.85%	0.97853	0.97942	0.97855

Among the top-3000 features, maximum number of features are selected from the theta frequency band as seen in Fig. 3. Fig. 4 plots the number of selected features associated with the 64 electrodes for the top-3000 selected features.

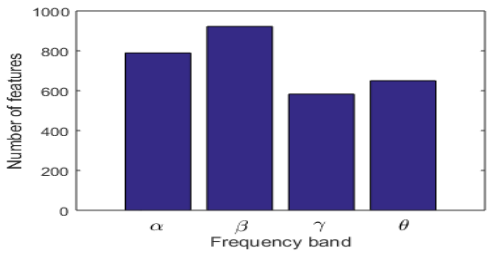


Fig. 3 Band-wise selected features

We observe from Fig. 4 that the maximum selected features are associated with the FP1, FP2, AF1 and AF2 electrodes; proving that the left and right frontal lobes are the brain regions that are affected the most in AUD as illustrated in Fig. 5. Our result demonstrates that the frontal lobe is the most vulnerable brain region in an alcoholic subject, a fact that is confirmed from the clinical study in [24].

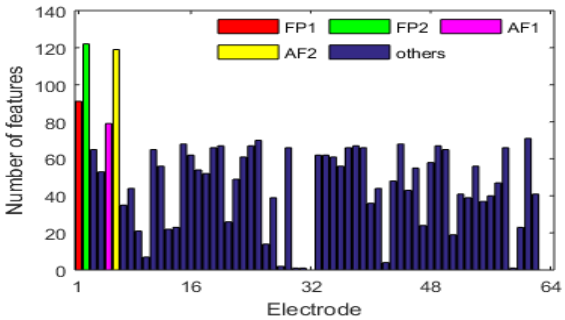


Fig. 4 Selected features from different electrodes

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:41:25 UTC from IEEE Xplore. Restrictions apply.

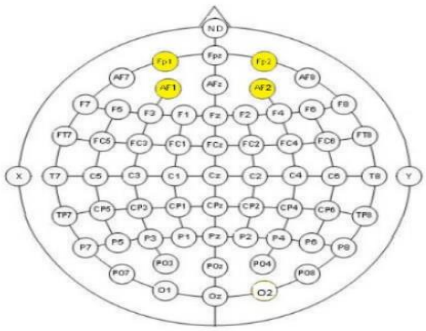


Fig. 5 The four electrodes associated with maximum number of selected features (FP1, FP2, AF1, and AF2) are shown highlighted in yellow.

V. CONCLUSION

In this work, we propose a novel QEEG feature set comprising of 7264 amplitude, spectral, range, and connectivity features, that are extracted from four frequency bands associated with the 64 EEG channels. The top-3000 features are voted by an ensemble of four state-of-the-art feature ranking algorithms and the resulting features are applied to four different classifiers for AUD detection. The results indicate that the reduced QEEG feature set achieves the highest accuracy of 97.85% with the ANN classifier, and outperforms the state of the art on QEEG classification. QEEG classification using spiking neural network models [23], that incorporate time information also, forms the future scope of our work. Another future work will be feature selection using feed-forward technique [25].

REFERENCES

[1] https://archive.ics.uci.edu/ml/machine-learning-databases/eeg-mld/

[2] GREENBLATT, MILTON and SIDNEY LEVIN. " The electroencephalogram associated with chronic alcoholism, alcoholic psychosis and alcoholic convulsions." Archives of Neurology & Psychiatry 52, no. 4 (1944): 290-295.

[3] Berry, M.S. and Pentreath V.W(1980): The neurophysiology of alcohol.In: Sandler, M(Ed.) Psychopharmacology of Alcohol. Raven press, New York, pp.43-72.

[4] Sharma, Manish, Dipankar Deb, and U. Rajendra Acharya. ”A novel three-band orthogonal wavelet filter bank method for an automated identification of alcoholic EEG signals.” Applied Intelligence 48, no. 5 (2018): 1368-1378.

[5] Bavkar, Sandeep, Brijesh Iyer, and Shankar Deosarkar. ”Optimal EEG channels selection for alcoholism screening using EMD domain statistical features and harmony search algorithm.” Biocybernetics and Biomedical Engineering 41, no. 1 (2021): 83-96.

[6] Mumtaz, Wajid, Pham Lam Vuong, Likun Xia, Aamir Saeed Malik, and Rusdi Bin Abd Rashid. ”An EEG-based machine learning method to screen alcohol use disorder.” Cognitive neurodynamics 11, no. 2 2017): 161-171.

[7] Shooshtari, M. Alavash, and S. Kamaledin Setarehdan. ”Selection of optimal EEG channels for classification of signals correlated with alcohol abusers.” In IEEE 10th INTERNATIONAL CONFERENCE ON SIGNAL PROCESSING PROCEEDINGS, pp. 1-4. IEEE, 2010.

[8] Anuragi, Arti, Dilip Singh Sisodia, and Ram Bilas Pachori. ”Automated alcoholism detection using fourier-bessel series expansion based empirical wavelet transform.” IEEE Sensors Journal 20, no. 9 (2020): 4914-4924.

[9] Bavkar, Sandeep, Brijesh Iyer, and Shankar Deosarkar. ”Rapid Screening of Alcoholism: An EEG Based Optimal Channel Selection Approach.” IEEE Access 7 (2019): 99670-99682.

[10] Fattah, S. A., K. Fatima, and C. Shahnaz. ”An approach for classifying alcoholic and non-alcoholic persons based on time domain features extracted from EEG signals.” In 2015 IEEE International WIE conference on Electrical and Computer Engineering (WIECON-ECE), pp. 479-482. IEEE, 2015.

[11] Bashiri, Azadeh, Leila Shahmoradi, Hamid Beigy, Behrouz A. Savareh, Masood Nosratabadi, Sharareh R. N Kalhori, and Marjan Ghazisaeedi. ”Quantitative EEG features selection in the classification of attention and response control in the children and adolescents with attention deficit hyperactivity disorder.” Future science OA 4, no. 5 (2018): FSO292.

[12] Upadhyay, R., P. K. Padhy, and P. K. Kankar. ”Alcoholism diagnosis from EEG signals using continuous wavelet transform.” In 2014 annual IEEE India Conference (INDICON), pp. 1-5. IEEE, 2014.

[13] Priya, Anchala, Pooja Yadav, Shweta Jain, and Varun Bajaj. ”Efficient method for classification of alcoholic and normal EEG signals using EMD.” The Journal of Engineering 2018, no. 3 (2018): 166-172.

[14] Ong, Kok-Meng, K-H. Thung, Chong-Yaw Wee, and Raveendran Paramesran. ”Selection of a subset of EEG channels using PCA to classify alcoholics and non-alcoholics.” In 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, pp. 4195-4198. IEEE, 2006.

[15] Mumtaz, Wajid, Pham Lam Vuong, Likun Xia, Aamir Saeed Malik, and Rusdi Bin Abd Rashid. ”Automatic diagnosis of alcohol use disorder using EEG features.” Knowledge-Based Systems 105 (2016): 48-59.

[16] Toole, John MO, and Geraldine B. Boylan. ”NEURAL: quantitative features for newborn EEG using Matlab.” arXiv preprint arXiv:1704.05694 (2017).

[17] Huang, Huijin, Zikang Niu, Gang Liu, Mengdi Jiang, Qingxia Jia, Xiaoli Li, and Yingying Su. ”Early consciousness disorder in acute large hemispheric infarction: an analysis based on quantitative EEG and brain network characteristics.” Neurocritical Care 33, no. 2 (2020): 376-388.

[18] Siuly, S., Bajaj, V., Sengur, A. and Zhang, Y., 2019. An advanced analysis system for identifying alcoholic brain state through EEG signals. International Journal of Automation and Computing, 16(6), pp.737-747.

[19] Zhang, Xiao Lei, Henri Begleiter, Bernice Porjesz, and Ann Litke. ”Electrophysiological evidence of memory impairment in alcoholic patients.” Biological Psychiatry 42, no. 12 (1997): 1157-1171.

[20] Shri, TK Padma, and Natarajan Sriraam. ”Pattern recognition of spectral entropy features for detection of alcoholic and control visual ERP’s in multichannel EEGs.” Brain informatics 4, no. 2 (2017): 147-158.

[21] O’Reilly, Deirdre, Michael A. Navakatikyan, Marcia Filip, Deirdre Greene, and Linda J. Van Marter. ”Peak-to-peak amplitude in neonatal brain monitoring of premature infants.” Clinical neurophysiology 123, no. 11 (2012): 2139-2153.

[22] Susan, Seba, and Madasu Hanmandlu. ”Smaller feature subset selection for real-world datasets using a new mutual information with gaussian gain.” Multidimensional Systems and Signal Processing 30, no. 3 (2019): 1469-1488.

[23] Holker, Ruchi, and Seba Susan. ”Neuroscience-Inspired Parameter Selection of Spiking Neuron Using Hodgkin Huxley Model.” InternationalJournal of Software Science and Computational Intelligence (IJSSCI)13, no. 2 (2021): 89-106

[24] Oscar-Berman, Marlene, and Ksenija Marinkovic. ”Alcoholism and thebrain: an overview.” Alcohol research health 27, no. 2 (2003): 125.

[25] Joadder, Md A. Mannan, Joshua J. Myszewski, Mohammad H. Rahman,and Inga Wang. ”A performance based feature selection technique forsubject independent MI based BCI.” Health information science andsystems 7, no. 1 (2019): 1-10.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:41:25 UTC from IEEE Xplore. Restrictions apply.

Silicon
https://doi.org/10.1007/s12633-021-01288-w

REVIEW PAPER

Recent Technological Advancement in Surrounding Gate MOSFET for Biosensing Applications - a Synoptic Study

Amit Das[1] · Sonam Rewari[2] · Binod Kumar Kanaujia[1] · R. S. Gupta[3]

Received: 26 April 2021 / Accepted: 20 July 2021
© Springer Nature B.V. 2021

Abstract
Advancement in the technology and device engineering has a huge impact on the evolution of biosensors and their designing principles. Various biosensors have been designed and reported depending upon the application it will be serving. Different biosensors have played an important role in vital applications in multifarious fields such as medical field, food industry, ecological monitoring and so on. The synoptic review chronicles the basic concepts, working principle, and structure of biosensors the surrounding gate MOSFET based biosensors. This paper presents a brief synoptic review about the generalization of the Surrounding Gate MOSFET based biosensors and their recent technological advancements. The recent advancements done in the device structure and material of the surrounding gate MOSFET based biosensors to improve sensitivity and to serve different sensing applications has been discussed in detail.

Keywords Biosensors · Charge plasma technique · Junctionless transistors · Schottky barrier · Surrounding Gate MOSFET biosensors

1 Introduction

A biosensor [1] is a powerful sensing analytical device in the field of bioelectronics [2, 3] that has merged two different major fields of biotechnology and electronics engineering. They have been evolving continuously from the time its idea came into the picture and has been always a hot topic of research [4] among the researchers all over the world. Efficiently designed biosensors [5] have been fabricated to sense the presence of biomolecules in the given sample which is currently trending in the medical research area [6]. The diurnal usage of the various electronic sensors [7] to measure the different biological system parameters has completely transformed the present-day medical diagnosis area [8–10].The size of biomolecule may vary from several micrometers to hundred of femto-metres [11]. Biosensors can be broadly categorized depending upon the transducers and biological receptors. Biosensors which are tissue based [12–16], enzyme based [17], affinity based [10], DNA based [18–22], antibody based [23–26], cell based [27–30], etc. falls under the category of biosensors categorized on the basis of biological receptor. Biosensors such as optical biosensors [31, 32], magnetic biosensors [33–35], electro chemical biosensors [36], calorimetric biosensors [37, 38] and mass based biosensors falls under the category of biosensors categorized on the basis of transducer. Here the topic of interest is surrounding gate MOSFET biosensor which is also called surrounding gate metal oxide semiconductor field-effect-biosensor (FEB). It has been a hot potential topic for research in both the industrial and academic domain. Any good biosensor [39] must be able to output and sense the minute changes arising because of the biological response. That means a good biosensor must be

✉ Sonam Rewari
rewarisonam@gmail.com

Amit Das
amitofficial7492@gmail.com

Binod Kumar Kanaujia
bkkanaujia@yahoo.co.in

R. S. Gupta
rsgupta1943@gmail.com

[1] School of Computational and Integrative Sciences, Jawaharlal Nehru University, 110067 New Delhi, India

[2] Department of Electronics and Communication Engineering, Delhi Technological University, 110042 New Delhi, India

[3] Department of Electronics and Communication Engineering, Maharaja Agrasen Institute Of Technology, 110086 New Delhi, India

Published online: 20 August 2021

highly efficient and sensitive in converting a biological response and output that response in the form of electrical signal.

A biosensor, in general consists of a device which will perform the biosensing operation. Figure 1 shows the basic schematic representation of a biosensor. The basic advantage of a surrounding gate MOSFET biosensor over the conventional MOSFET biosensors is that the former one has a better electrostatic control over the charge carriers flowing across the channel. Surrounding gate MOSFET is a simple yet valuable structural variant of MOSFET. It is very much similar to a FinFET in its construction technique except the gate. The gate is surrounding all over the channel and now the gate has better control over the channel and the movement of the electrons through the channel. And since the gate is surrounding all over the channel all, this variant is also known as Gate-All-Around-MOSFET. The first ever surrounding gate MOSFET device was exemplified and demonstrated by Toshiba and the research team was led by Kazumasa Sunouchi, Hiroshi Takato and Fujio Masuoka [40]. With the extensive research going on, many modified variants of surrounding gate MOSFET were reported in the recent years. Basic structure of surrounding gate MOSFET [41–46] was modified to incorporate multiple gates, multiple gate oxide layers, core shell architecture, underlap and overlap gate architecture were reported in the last few years. Each variant [47–51] came up with its own advantages and shortcomings. Material engineering was also incorporated with the structural engineering such as using high k dielectric material for gate oxide, gate oxide stacking technique with two or more layers of gate oxide, using InGaAs or Ge as a substitute of silicon. A good biosensor should be reusable, independent of operating conditions and physical parameters and should be highly specific in detecting the particular analyte.

2 Operating Principle Behind the Sensing Ability of Surrounding Gate MOSFET

Basic operation of a surrounding gate MOSFET is similar to a conventional MOSFET device except that of the gate coverage area surrounding the channel which will affect the flow of electrons flowing across the channel. Biosensing [39] ability of MOSFET is one of the important application that it can competently perform. The charge carriers start moving from the source terminal under the influence of voltage applied across the source and drain terminal. The voltage applied across the gate influences the flow of the charge carriers. The sensing ability of any surrounding gate MOSFET refers to the device metrics that will linearly change with the analyte's concentration and its charge which can be measured through external interfacing circuitry. There are many device metrics that will change with the concentration and charge of the analyte. Device metrics such as drain current, threshold voltage, subthreshold slope, on-current to off-current ratio, transconductance, output conductance etc. are useful while testing the sensitivity of our surrounding gate MOSFET.

The introduction of an analyte influences the output quantities of any device. Since it is known that the flow of charge carriers is being influenced by the gate and the gate oxide layer parameters. So if we are able to incorporate the analyte (which has to be measured) within our device physical area, then the analyte along with the gate will influence the flow of charge carriers. The best place to incorporate the analyte to be tested will be near the gate oxide. It means that we have to incorporate the analyte by creating a cavity in the gate oxide and this is the key to basic operation behind a surrounding gate MOSFET. The charge and the concentration of the analyte will influence the potential across the channel. The flow of the charge carriers will be influenced by the modified electric field which will simultaneously change the various output quantities. This change in the output quantity will work as the sensing metric for the device and the aim of any sensing device is to keep the value of sensitivity as high as possible. Figure 2 shows the generalized biosensing principle of a surrounding gate MOSFET biosensor.

3 Background and Progress

This section deals with the background and generalization of the work done in the field of surrounding gate MOSFET biosensors. Basically the work that has been done in the field of surrounding gate biosensors can be broadly classified into:

1) Simulation based study,
2) Analytical modeling based study,
3) Comparative analysis based study.

But the thing that is common in all these papers is that every paper had proposed a different variant of surrounding

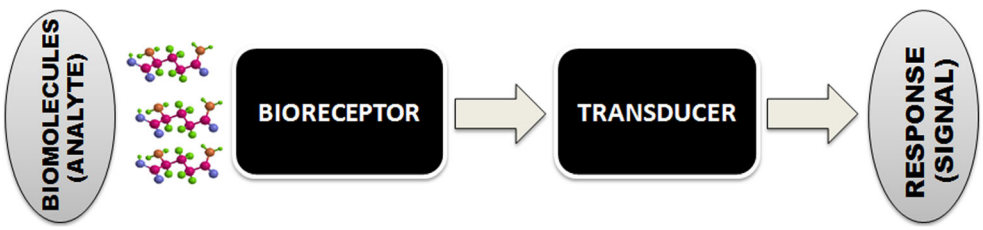


Fig. 1 A generalized schematic representation of a biosensor

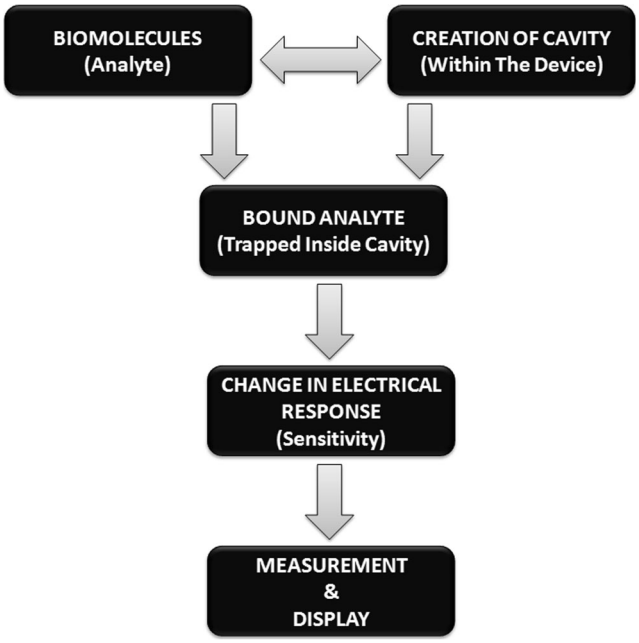
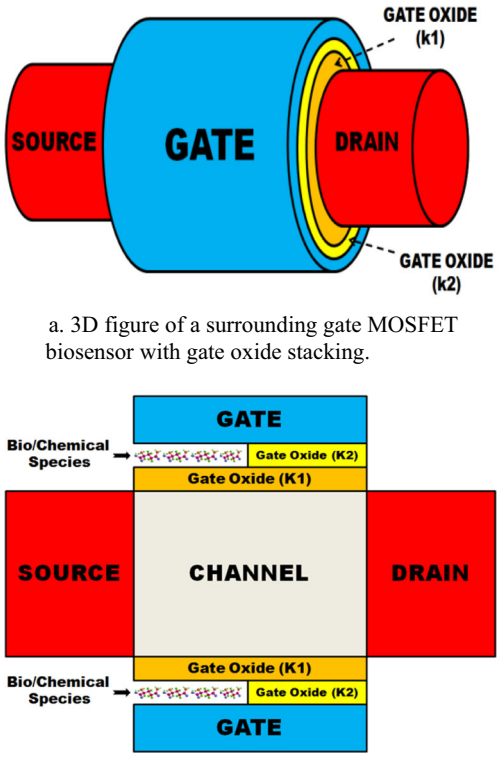


Fig. 2 Basic principle of a surrounding gate MOSFEB



a. 3D figure of a surrounding gate MOSFET biosensor with gate oxide stacking.

b. Cross-sectional view of a surrounding gate MOSFET biosensor with gate oxide stacking.

Fig. 3 **a** 3D figure of a surrounding gate MOSFET biosensor with gate oxide stacking. **b** Cross-sectional view of a surrounding gate MOSFET biosensor with gate oxide stacking

gate MOSFET which can be used for biosensing application, so this particular section deals with the generalization of all those structures of surrounding gate MOSFET biosensors that have been reported till now. All the structures or variants of surrounding gate MOSFET biosensors can be broadly classified into few sub-categories which have been discussed briefly below.

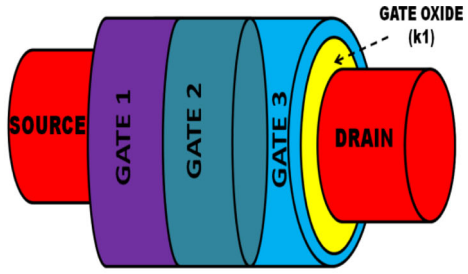
3.1 Surrounding Gate MOSFET Biosensor With Gate Oxide Stacking

This is one of the variant of the surrounding gate MOSFET biosensor which utilizes the gate oxide stacking technique [11]. Gate oxide stacking refers to the use of more than one layer of gate oxide. Few structures use stacking of two layers of high-k and low-k gate oxide materials whereas a few use more than two layers of gate oxide of different dielectric constant value. The basic principle modification done in this structure is creating a cavity which will trap the biomolecules to serve the function of a biosensor.

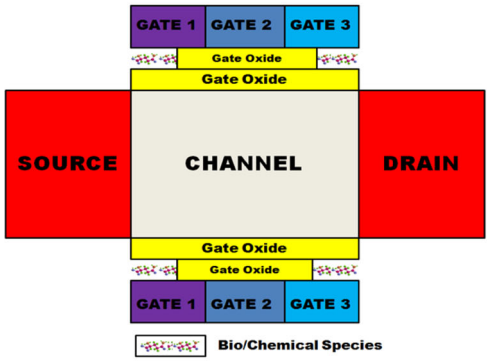
A typical 3D surrounding gate MOSFET biosensor with the gate oxide stacking and its cross sectional view is shown in the Fig. 3(a) and (b). Here, even along with the choice of semiconductor material, the doping level of the source/channel/drain, work-function of the gate along with the placement and the dimension of the cavity which affects the sensitivity of the surrounding gate MOSFET biosensor, the sensitivity is also affected by the number of gate oxide layers and their placement order. The position, length and thickness of the cavity will differ in different structures of biosensors which will ultimately affect the sensitivity and different device sensing metrics of the biosensor. Biomolecules are immobilized inside the cavity. Immobilization of the biomolecules inside the cavity leads to the change of electric field and gate oxide capacitance which ultimately changes the device output parameters. Generally, the gate oxide stacking has been commonly used along with the other variants of the surrounding gate MOSFET structure.

3.2 Surrounding Multi-Gate MOSFET Biosensor

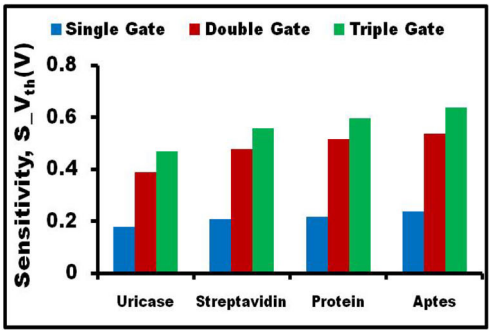
This particular section deals with the multiple gate variants of surrounding gate MOSFET biosensor [43, 52]. Surrounding gate MOSFET biosensors with two and three gates has been reported so far but with the planar MOSFET structures, a four gate structure has been reported too. These gates might operate with the common gate voltage or with different biasing voltage. Literature review reveals that the work-function value of the gate also affects the sensitivity of the device. Figure 4(a) and (b) shows 3D structure and cross sectional view of a generalized surrounding triple gate MOSFET biosensor. Even structures with more than two gates have been reported till now. Although the performance gets better with the multiple gates structure but the fabrication complexity also keeps on increasing with the increasing number of gate since the



a. 3D figure of a surrounding double gate MOSFET biosensor.



b. Cross sectional view of a surrounding triple gate MOSFET biosensor.



c. Sensitivity pattern due to multi-gate surrounding gate MOSFET biosensor.

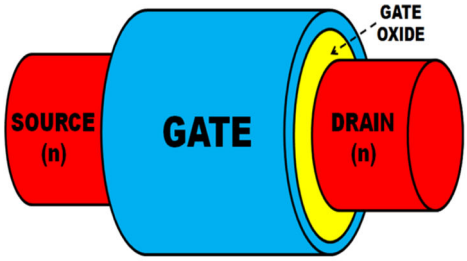
Fig. 4 **a** 3D figure of a surrounding double gate MOSFET biosensor. b Cross sectional view of a surrounding triple gate MOSFET biosensor. **c** Sensitivity pattern due to multi-gate surrounding gate MOSFET biosensor

fabrications is being done on a nanometer scale. Here, the sensitivity is affected by the number of gates, placement of gate and the work-function value of the gates used in the fabrication of biosensor. Anubha Goel [52] discussed a triple gate MOSFET biosensor in 2016 and demonstrated an improved sensitivity pattern over single or dual gate MOSFET biosensors which has been shown in the Fig. 4(c). It was shown and demonstrated in the paper that sensitivity improves by a significant amount in a triple gate structure when compared to sensitivity offered by a single or double gate structure. In general, multiple gates with different work-function (Φ) offers a better control over the charge carriers flowing across the channel which again affects the sensitivity pattern of the device.

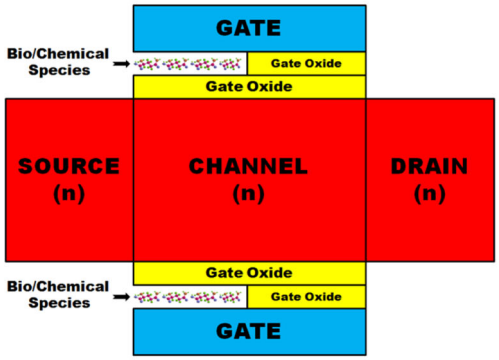
3.3 Junctionless Surrounding Gate MOSFET Biosensor

This particular section deals with the junctionless surrounding gate MOSFET based biosensors [11]. Scaling down of devices to accommodate more transistors on a single chip has led to shortening of channel and with a short channel device, junction formed between channel-source and channel-drain can led to some undesired results with a short effective channel length than the actual length. In the fabrication of a junctionless surrounding gate MOSFET biosensor, a single bar of same material is used for creation of channel, source and drain (doping remains same for source/channel/drain). Generally a semiconductor bar of high doping is always preferred. Also a junctionless transistor is more immune to the various short channel effects when compared to a transistor working in inversion mode [11]. Since there is a uniform doping throughout from the source to the drain terminal which also includes the channel, so the working transistor principle of a junctionless surrounding gate MOSFET will be different from the inversion mode surrounding gate MOSFET biosensors. The main thing is that the biosensing principle will remain almost same in both the cases with the former device showing an improvement in biosensing action. The difference in the work-function of the channel and the gate is used for depleting the charge carriers in the channel. The main advantage of this junctionless surrounding gate MOSFET biosensor is the fabrication simplicity over the other variants since we don't have to fabricate the source and drain part separately for the same device.

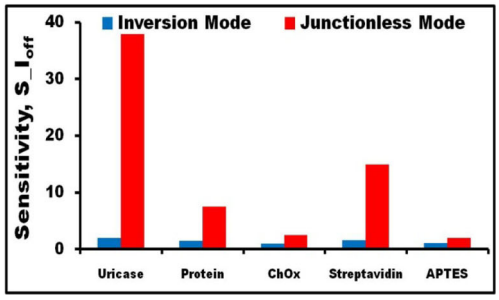
Figure 5(a) and (b) shows the 3D and cross sectional view of a generalized junctionless surrounding gate biosensor. Again, this sensitivity is highly dependent upon the choice of selection of the type of the material used and its doping concentration along with the placement of the cavity and its dimension. Junctionless refers to the absence of any junction formation between the source-channel and drain-source interface because a single uniform bar of semiconductor is used as the fabricating material for the source, channel and drain. The doping of the bar is kept high in the fabrication of junctionless devices so as to ensure that a large number of carriers are supplied by the source. Due to the absence of junction, these biosensors show a better suppression of short channel effects when compared to the inversion-mode biosensors. These biosensors show a comparable high sensitivity because of high doping and bulk conduction since inversion mode biosensors show surface conduction. Yogesh Pratap and Manoj Kumar [11] presented a paper in 2017 that was purely based upon the concept of junctionless surrounding gate FET biosensors. They demonstrated the use of this device for the label-free detection of ChOx, Protein, APTES, Streptavidin and



a. 3D figure of a junctionless surrounding gate MOSFET biosensor.



b. Cross-sectional view of a junctionless surrounding gate MOSFET biosensor.



c. Sensitivity comparison in junctionless surrounding gate MOSFET biosensor

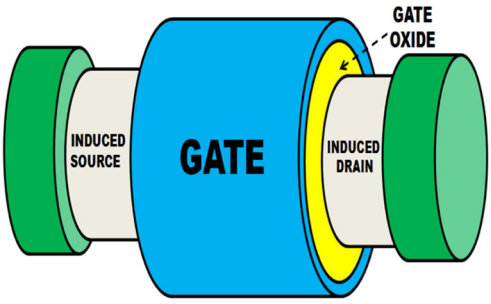
Fig. 5 **a** 3D figure of a junctionless surrounding gate MOSFET biosensor. **b** Cross-sectional view of a junctionless surrounding gate MOSFET biosensor. **c** Sensitivity comparison in junctionless surrounding gate MOSFET biosensor

Uricase (neutral biomolecules). They had compared and demonstrated the improved sensitivity of the proposed device over the inversion mode GAA MOSFET based biosensors. It was also epitomized in the paper that the drain current or threshold voltage based sensing was inferior when compared to the capacitance based sensing. Also, it was pointed in the paper that if the proposed device is to be operated in the ohmic region, then the capacitance based sensing is much more effective but if the same device is to be operated in the sub-threshold region, and then the drain current based sensing is much more effective. Figure 5(c) shows the comparison of sensitivity value of off-current in junctionless and inversion mode for a surrounding gate MOSFET biosensor.

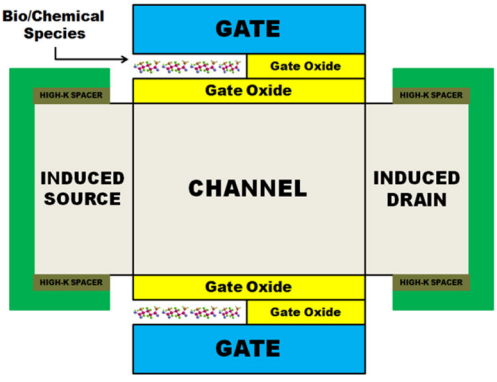
3.4 Charge Plasma based Surrounding Gate MOSFET Biosensor

This particular section deals with the charge plasma surrounding gate MOSFET based biosensors [53]. Similar to the junctionless surrounding gate MOSFET biosensors, scaling down of device feature size to fit-in more transistors on a single chip has led to shortening of channel and with a short channel device, junction formed between channel-source and channel-drain can led to some undesired results with a short effective channel length than the actual length. In the fabrication of charge plasma surrounding gate MOSFET biosensor, a single intrinsic bar of same semiconductor material is used for the fabrication of channel, source and drain part. The concept of charge plasma is used to create the source and drain terminals in the device. In the charge plasma technique [54], we use the metal of some work-function which will induced the n^+ or p^+ regions. Induction [55] of n^+ or p^+ region depends upon the work-function of metal used. It means the source and the drain terminal used here are the induced source and drain.

Figure 6(a) and (b) shows the generalized 3D structure and the cross sectional view of a charge plasma surrounding gate MOSFET biosensor. A p^+ or n^+ type induction will purely



a. 3D figure of a charge-plasma based surrounding gate MOSFET biosensor.



b. Cross-sectional view of a charge-plasma based surrounding gate MOSFET biosensor.

Fig. 6 **a** 3D figure of a charge-plasma based surrounding gate MOSFET biosensor. **b** Cross-sectional view of a charge-plasma based surrounding gate MOSFET biosensor

depends upon the work-function of the metal used around the source and the drain terminal. If $\Phi_M < \chi_{Ch} + E_g/2$, a n type induction occurs in the semiconductor bar but if $\Phi_M > \chi_{Ch} + E_g/2$,then a p type induction occurs in the semiconductor bar(where symbols denotes the usual meaning, Φ_M : work-function of the metal, χ_{Ch} : electron affinity, E_g : energy band-gap of semiconductor material). We don't have to create diffused source/drain, instead the source/drain is induced which eliminates one to two steps of fabrication process, thus making the fabrication of these biosensors comparatively easier. Rahul Das and Manash Chandra [53] demonstrated analytically the sensitivity pattern of surrounding gate MOSFET based biosensor based upon the charge plasma based concept. The threshold voltage and I_{ON}/I_{OFF} had been used as the device metrics here which had shown an improved electrostatics and sensitivity with a very much simplified fabrication process, reduced random dopant fluctuations and easier control of the doping which was investigated for the both charged and uncharged biomolecules.

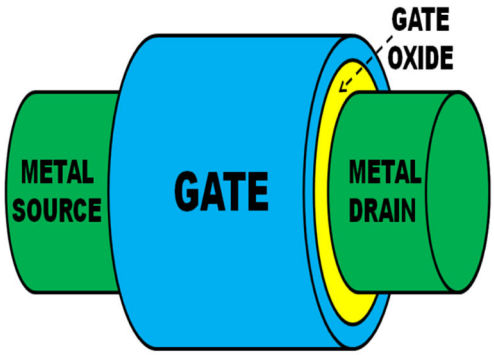
3.5 Material Engineered Surrounding Gate MOSFET Biosensor

This section deals with the material engineered surrounding gate MOSFET biosensor. Material engineering refers to the usage of different materials for the fabrication of any device. We may use elemental semiconductor (Si or Ge), binary compound semiconductor(II-IV compounds, III-V compounds, IV-IV compounds), ternary compound semiconductor(GaAsP), quaternary compound semiconductor (InGaAsP) material in the fabrication of surrounding gate MOSFET biosensor [42]. Although a much material engineering has been explored in the MOSFET variants and a few papers have utilized the concept of material engineering in the planar variants of MOSFET biosensors. Material engineering does not only refer to the usage of different semiconductor materials but it also covers the usage of different oxide materials like HfO_2, $Al_2O_3$and ZnO, and it also includes the usage of possible alternatives of metallic gate such as poly-silicon gate or catalytic metal gates. Few researchers have used the poly-silicon gate instead of metallic gate in the fabrication of biosensors. Using poly-silicon gate instead of metallic gate has its own advantages in the fabrication process and this engineering technique comes under the section of gate engineering techniques. A very few papers are available which has utilized the concept of material engineering in the fabrication of surrounding gate MOSFET biosensors. Navaneet Kumar Singh and Ashish Raman [42] have presented a paper which utilizes a highly doped Indium-Gallium-Arsenide ($In_{1-x}Ga_xAs$) channel material for the surrounding gate nano-wire FET for sensing applications. This proved to produce and attain a high sensitivity for hydrogen and oxygen gas sensing applications when compared to the silicon nano-wire .Various device metrics such as on-current, off-current and on-current/off-current ratio have been used in the paper and effect of radius/channel have been studied to study the impact of these parameters on sensing ability of the proposed biosensor. The authors pointed out that decreasing radius or increasing channel length improves the sensitivity of the proposed device in that particular proposed device.

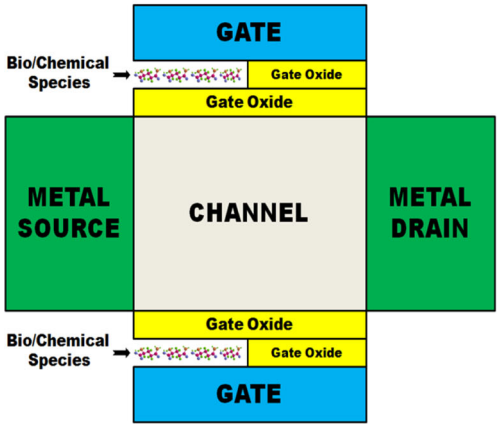
3.6 Surrounding Gate MOSFET Biosensor with Schottky Barrier

This section deals with the variants of surrounding gate MOSFET biosensors having Schottky barrier [41, 56]. Some of the surrounding gate MOSFET biosensors have been also designed to detect the various gases. A Schottky barrier is formed between a metal and a semiconductor material. The rectifying or the ohmic property of the Schottky barrier is determined by the Schottky barrier height which depends on both the metal and the semiconductor. Using a metal source and drain instead of doped semiconductor has the advantage of low parasitic resistance between the source-channel and drain-channel interface and fabrication advantage in the form of low temperature fabrication process of source and drain terminals. This Schottky barrier lowers the leakage flow in the device, high integration of different engineered materials because of low thermal budget processing [57].

Figure 7(a) and (b) shows the 3D and cross sectional view of a generalized surrounding gate MOSFET biosensor with Schottky barrier. In these types of biosensors, a metallic source and a metallic drain is used instead of using a source/drain fabricated from semiconductor material. A metallic drain is optional but metallic source is compulsory in the fabrication process. Now the junction is formed between a metal and a semiconductor material, so the barrier formed is Schottky barrier. This Schottky barrier in these biosensors uses the advantage of low resistivity when compared to the conventional junction formed between semiconductor materials. This low resistivity implies low resistance which means on-current (I_{ON}) is higher in such devices. Author Manoj Kumar [41] reported a dopant segregated surrounding gate MOSFET biosensor with Schottky barrier in 2016. The author had discussed that how the concept of dopant segregation can further improves the sensitivity of biosensor. He had taken shift in the current and threshold voltage as sensing metric for the proposed biosensor. CMOS compatibility, low power consumption, array type integration on a lab-on-a-chip are few of the advantages of the proposed structure. He pointed out in the paper that the proposed structure shows a high sensitivity towards low concentration negatively charged biomolecules. The high sensitivity is a result due to the bending of the bands at the drain/source end which ultimately increases tunneling probability.



a. 3D figure of a surrounding gate MOSFET biosensor with Schottky barrier.



b. Cross sectional view of a surrounding gate MOSFET biosensor with Schottky barrier.

Fig. 7 **a** 3D figure of a surrounding gate MOSFET biosensor with Schottky barrier. **b** Cross sectional view of a surrounding gate MOSFET biosensor with Schottky barrier.

3.7 Nanotube Based Surrounding Gate MOSFET Biosensor

This section deals with the variants of the surrounding gate MOSFET biosensors which utilizes the nanotube as the channel material. Nanotubes such as carbon nanotube, silicon nanotube etc. have been a popular channel material in the recent decade. Carbon nanotube is quite famous in the planar structures of MOSFET based biosensors but in the case of surrounding gate MOSFET biosensors, a very few papers have been presented so far with nanotube as the channel material. Carbon nanotube is just a regular cylindrical structure made up of carbon atoms. It is typically a single layered atom of carbon rolled-up to form a cylindrical structure and silicon nanotube is similar as it consists of silicon atoms. A gas sensor based upon silicon nanotube MOSFET sensor was presented by Ali Shahhoseini, Farad Assareh Poura in 2014 [45]. The authors pointed out that the proposed device increases the sensitivity and reduces consumption of power. This was mainly due to the increase in the sensing area surface. The main advantage of nanotube is that these are extremely flexible, smaller in size and light in weight but sometimes this advantage proves to be a disadvantage in the fabrication process of surrounding gate MOSFET biosensors. That's why, a very few papers have been reported till now which uses nanotube in the fabrication of surrounding gate MOSFET biosensors.

3.8 Hybrid Surrounding Gate MOSFET Biosensor

This section represents the hybrid structures of surrounding gate MOSFET biosensors. The basic variants discussed in the above sections have their own advantage and disadvantage. So, generally a combination [43] of these different variants [58] is used to fabricate a surrounding gate biosensor. In most of the papers, gate oxide stacking has been frequently used with other variants The papers reviewed till now have shown that few variants like junctionless surrounding gate MOSFET biosensor shows better suppression of short channel effects while other variants shows improvement in terms of sensitivity. Even these biosensors shows varied sensitivity towards charged and neutral biomolecules, so particularly the selection (combination of variants) will also depend upon the type of application for which biosensor has to be fabricated.

Table 1 shows the comparison of main variants of surrounding gate MOSFET based biosensors in a tabulated form where sensitivity has been compared for those variants. All the above variants are of surrounding gate MOSFET which have been used for sensing applications. The table shows the variant type, parameter considered for sensitivity analysis, sensitivity value at different operating conditions and modeling done for the analysis of the variant. Although structure might be different of all the variants but this table roughly gives the idea about how much effective the variant is for the sensing applications. By seeing the table, it's evident that the junctionless and material engineered biosensors are quite sensitive which is desirable for the sensing applications. The table summarizes the papers that are used in the literature review. Different authors have used different operating conditions to simulate the proposed biosensor and some of them have analytically verified their results. Even different authors have used different sensing parameters to test the sensitivity of their device. FET based devices are not only suitable for the biosensing applications but can be used for other applications also [59].

4 Practical Challenges

Although surrounding gate MOSFET biosensors have proven to be worthy in day to day life but still, there are various

Table 1 Comparison of different variants of surrounding gate MOSFET biosensors

Variants of surrounding gate MOSFET biosensors																				
	Surrounding Multiple-Gate MOSFET Biosensor [52]		Junction-less Surrounding Gate MOSFET Biosensor [11]		Charge Plasma Based Surrounding Gate MOSFET Biosensor [53]				Material Engineered Surrounding Gate MOSFET Biosensor [42]				Surrounding Gate MOSFET Biosensor With Schottky Barrier [41]				Nanotube Based Surrounding Gate MOSFET Biosensor [45]			
Variant	Dielectric Modulated Triple Gate Type		Junction-less Type		Charge Plasma Type				Material Engineering ($In_{1-x}Ga_xAs$) Type				Schottky Barrier Type				Nanotube Channel Type			
Parameter considered	Threshold Voltage S_{V_t}		Drain OFF Current $S_{I_{OFF}}$		A=Threshold Voltage B=Drain ON Current ρ_1=-0.4e12 C/cm^2 ρ_2=-0.2e12 C/cm^2 ρ_3=0 ρ_4=0.2e12 C/cm^2 ρ_5=0.4e12 C/cm^2				A=Off current $S_{I_{OFF}}$ B=On current $S_{I_{ON}}$				Threshold Voltage A=ΔV_{th-SB} B=ΔV_{th-DS_SB} L=40nm R=10nm T_{N-Gap}=1nm $L_{seg/D}$=$L_{seg/S}$=5nm				OFF Current A=Oxygen Sensing B=Hydrogen Sensing			
Sensitivity value	Difference		Fraction		Difference				Fraction				Fraction				Fraction			
	A:Single Gate		R		A		B		Φ	A	B			A	B			A	B	
	B:Dual Gate		(10nm)				10^{-5}				10^{-3}									
	C:Triple Gate		L																	
	URICASE		(50 μm)		K	K	K	K	H_2 Gas Sensing				K	0.105	0.19		ΔΦ=100mev			
	A: 0.18		Cavity Width		12	2	12	2					3							
	B: 0.39		(4nm)		ρ_1	ρ_1	ρ_1	ρ_1	50	6.3	33.89		K	0.13	0.23		Oxide Thickness=2nm			
	C:0.47				0.13	0.18	4.9	0	mev				5				Device Length=150nm			
																	Channel Length=50nm			
																	AIR Core			
	STREPT-AVIDIN								100	39.51	76.27						R			
	A: 0.21								mev								nm			
	B:0.48				ρ_2	ρ_2	ρ_2	ρ_2	150	245.45	110.16		K	0.21	0.25		R	31.1	33.6	
	C:56				0.08	0.08	4.95	1.2	mev				7				5			
	PROTEIN		2						200	1509.09	152.54									
	A: 0.22		APTES						mev											
	B:0.52		38		ρ_3	ρ_3	ρ_3	ρ_3	O_2 Gas Sensing				K	0.16	0.26		R	23.6	26.8	
	C:0.6		URICASE		0.02	0.025	5	1.5					10				7			
	APTES		15		ρ_4	ρ_4	ρ_4	ρ_4	50	6.33	37.03		ρ	1.02	0.62		R	15.2	18.1	
	A: 0.24		STREPT-AVIDIN		0.01	0.061	5.04	2	mev				-				10			
	B:0.54								100	40.60	76.92		3e12							
	C:0.64								mev											
			7.5		ρ_5	ρ_5	ρ_5	ρ_5	150	257.37	116.23		ρ	1.04	1.04		R_{in}	42	38.7	
			PROTEIN		0.02	0.10	5.1	2.2	mev				+				7			
			2.5						200	1688.17	159.544		3e12				&			
			ChOx						mev								R_o			
																	10			
Modeling	Simulation		Analytical & Simulation		Analytical & Simulation				Simulation				Simulation				Simulation			

challenges that may arise and could be a hindrance in the sensing action of the surrounding gate MOSFET biosensors. All the proposed devices are simulated in a simulator which assumes ideal conditions but there are practical challenges that could surely affect the results in non-ideal conditions.

Fabrication Challenge Creating a precise cavity of the mentioned dimension in a nano-scale device.

Selectivity The designed device must be sensitive to the specific analyte. This means that the biosensor must respond only to the targeted biomolecule but in reality the biosensor responds to few non targeted biomolecule.

Sensitivity Sensitivity refers to the change in the device output quantity with the introduction of the target biomolecule. This must be as high as possible. Although, most of the biosensors reported so far have shown a high sensitivity for neutral biomolecules and negatively charged biomolecules but the sensitivity is quite low for the positively charged biomolecules. A highly sensitive biosensor must be highly sensitive to detect even a single biomolecule.

Short Channel Effects Scaling down of the device physical size gives rise to various short channel effects like velocity saturation, impact ionization, surface scattering, drain induced barrier lowering, hot carrier injection. Some of the engineering techniques like gate engineering or junctionless can reduce the effects of short channel effects but even they are unable to fully eradicate these effects.

Reusability Reusability of the biosensors is related to the economic concern. All these biosensors are susceptible to wear and tear damage because of the physical interactions between the biomolecule and the biosensor.

Filling Area This refers to the filling up of the cavity formed inside a biosensor. Since, the performance of the biosensor is highly affected by the cavity dimensions and the sensitivity, therefore a fully filled cavity ensures highly efficient sensing action by the biosensor. Designing a cavity of any shape may be easy to form but ensuring a cent percent filled area of the cavity is a concern which needs to be solved.

Signal to Noise Ratio This is defined as the ratio of actuating signal that is electrically transduced from the targeted biomolecules to the non-actuating signals that are electrically transduced from the non-targeted biomolecules. This signal to noise ratio is one of the important parameter to judge the efficiency of any biosensor and it will be high if the sensitivity along with the selectivity of the biosensor is high.

5 Conclusions

This paper presents a précis of the generalization of the latest work done in the structural variation of surrounding gate MOSFET based biosensors. The usefulness of these surrounding gate MOSFET based biosensor can be confirmed with its increased usability in our daily lives. Sensitivity of different reviewed biosensors highly depends upon the cavity dimensions and placement. The length, thickness and even the position of the cavity can highly affects the sensitivity of the device. Some of the variants like junction-less surrounding gate MOSFET biosensors were able to suppress the short channel effects better than the others and few of the structures show a tremendous improvement in terms of sensitivity. The problem of filling area and selectivity towards the desired biomolecule in real time might deviate the practical output of any biosensor from the simulated output (desired one). Overall, the surrounding gate MOSFET shows a good potential in serving as an efficient nano-scale and highly sensitive biosensors to meet the future demands.

Acknowledgements Authors are grateful to the Director, Maharaja Agrasen Institute of Technology, Delhi for providing the facilities to carry out this research work. One of the author (Amit Das) would like to thank UGC, Government of India for financial assistance to carry out this research work.

Author Contributions The authors have contributed mutually regarding this paper.

Data Availability Not Applicable.

Declarations

Conflict of Interest There are no conflicts of interest amongst the authors.

Consent to Participate All the authors have complete consent to participate.

Consent for Publication All the authors have complete consent for publication.

References

1. Polizzi KM (2019) Biosensors. In: Comprehensive biotechnology
2. Karunakaran C, Bhargava K, Benjamin R (2015) Biosensors and bioelectronics. Elsevier Science, Amsterdam
3. Szent-Györgyi A (1968) Bioelectronics. Science 161(3845):988–990
4. Mehrotra P (2016) Biosensors and their applications - A review. J Oral Biol Craniofac Res 6(2):153–159
5. Koyun A, Ahlatcolu E, Koca Y (2012) Biosensors and their principles. In: A roadmap of biomedical engineers and milestones, no. 1, InTech, London, pp 114–142

6. Ramasamy R, Gopal N, Kuzhandaivelu V, Murugaiyan S (2014) Biosensors in clinical chemistry: An overview. Adv Biomed Res 3(1):67–75
7. Jianrong C, Yuqing M, Nongyue H, Xiaohua W, Sijiao L (2004) Nanotechnology and biosensors. Biotechnol Adv 22(7):505–518
8. Rivet C, Lee H, Hirsch A, Hamilton S, Lu H (2011) Microfluidics for medical diagnostics and biosensors. Chem Eng Sci 66(7):1490–1507
9. D'Orazio P (2003) Biosensors in clinical chemistry. Clin Chim Acta 334(1-2):41–69
10. Liu Y, Matharu Z, Howland MC, Revzin A, Simonian AL (2012) Affinity and enzyme-based biosensors: Recent advances and emerging applications in cell analysis and point-of-care testing. Anal Bioanal Chem 404(4):1181–1196
11. Pratap Y, Kumar M, Kabra S, Haldar S, Gupta RS, Gupta M (2018) Analytical modeling of gate-all-around junctionless transistor based biosensors for detection of neutral biomolecule species. J Comput Electron 17(1):288–296
12. Balint R, Cassidy NJ, Cartmell SH (2014) Conductive polymers: Towards a smart biomaterial for tissue engineering. Acta Biomater 10(6):2341–2353
13. De A, Jasani A, Arora R, Gambhir SS (2013) Evolution of BRET biosensors from live cell to tissue-scale in vivo imaging. Front Endocrinol (Lausanne) 4:1–6
14. Campàs M, Carpentier R, Rouillon R (2008) Plant tissue-and photosynthesis-based biosensors. Biotechnol Adv 26(4):370–378
15. Hasan A et al. (2014) Recent advances in application of biosensors in tissue engineering. Biomed Res Int 2014
16. Holmes TC (2002) Novel peptide-based biomaterial scaffolds for tissue engineering. Trends Biotechnol 20(1):16–21
17. Sassolas A, Blum LJ, Leca-Bouvier BD (2012) Immobilization strategies to develop enzymatic biosensors. Biotechnol Adv 30(3):489–511
18. Junhui Z, Hong C, Ruifu Y (1997) DNA based biosensors. Biotechnol Adv 15(1):43–58
19. Zhao WW, Xu JJ, Chen HY (2014) Photoelectrochemical DNA biosensors. Chem Rev 92(5):4094–4100
20. Pei H, Zuo X, Pan D, Shi J, Huang Q, Fan C (2013) Scaffolded biosensors with designed DNA nanostructures. NPG Asia Mater 5(6)
21. Wang J (2000) From DNA biosensors to gene chips. Nucleic Acids Res 28(16):3011–3016
22. Chao J, Zhu D, Zhang Y, Wang L, Fan C (2016) DNA nanotechnology-enabled biosensors. Biosens Bioelectron 76:68–79
23. Hock B, Seifert M, Kramer K (2002) Engineering receptors and antibodies for biosensors. Biosens Bioelectron 17(3):239–249
24. Sharma S, Byrne H, O'Kennedy RJ (2016) Antibodies and antibody-derived analytical biosensors. Essays Biochem 60(1):9–18
25. Zeng X, Shen Z, Mernaugh R (2012) Recombinant antibodies and their use in biosensors. Anal Bioanal Chem 402:3027–3038
26. North JR (1985) Immunosensors: Antibody-based biosensors. Trends Biotechnol 3(7):180–186
27. Ahmed A, Rushworth JV, Hirst NA, Millner PA (2014) Biosensors for whole-cell bacterial detection. Clin Microbiol Rev 27(3):631–646
28. Lagarde F, Jaffrezic-Renault N (2011) Cell-based electrochemical biosensors for water quality assessment. Anal Bioanal Chem 400(4):947–964. doi:https://doi.org/10.1007/s00216-011-4816-7
29. Liu Q, Wu C, Cai H, Hu N, Zhou J, Wang P (2014) Cell-based biosensors and their application in biomedicine. Chem Rev 114(12):6423–6461
30. Wang J, Wu C, Hu N, Zhou J, Du L, Wang P (2012) Microfabricated electrochemical cell-based biosensors for analysis of living cells in vitro. Biosensors 2(2):127–170
31. Monk DJ, Walt DR (2004) Optical fiber-based biosensors. Anal Bioanal Chem 379:931–945
32. Fan X, White IM, Shopova SI, Zhu H, Suter JD, Sun Y (2008) Sensitive optical biosensors for unlabeled targets: A review. Anal Chim Acta 620(1-2):8–26
33. Nabaei V, Chandrawati R, Heidari H (2018) Magnetic biosensors: Modelling and simulation. Biosens Bioelectron 103:69–86
34. Koh I, Josephson L (2009) Magnetic nanoparticle sensors. Sensors 9(10):8130–8145
35. Jamshaid T et al (2016) Magnetic particles: From preparation to lab-on-a-chip, biosensors, microsystems and microfluidics applications. TrAC - Trends Anal Chem 79:344–362
36. Wang J (2008) Electrochemical glucose biosensors. Chem Rev 108(2):814–825
37. Danielsson B (1991) Calorimetric biosensors. Biochem Soc Trans 19(1):26–28
38. Park SC et al. (2007) A calorimetric biosensor and its application for detecting a cancer cell with optical imaging. IFMBE Proc14(1): 637–640
39. Turner APF (2013) Biosensors: Sense and sensibility. Chem Soc Rev 42(8):3184–3196
40. Fahad HM, Hussain MM (2012) Are nanotube architectures more advantageous than nanowire architectures for field effect transistors? Sci Rep 2(1):2–8
41. Kumar M, Haldar S, Gupta M, Gupta RS (2016) DS Schottky barrier cylindrical GAA MOSFET: nanosensor for biochips. Nanomater Energy 5(1):10–19
42. Singh NK, Raman A, Singh S, Kumar N (2017) A novel high mobility In1-xGaxAs cylindrical-gate-nanowire FET for gas sensing application with enhanced sensitivity. Superlattices Microstruct 111(February 2019):518–528, [Online]. Available: https://doi.org/10.1016/j.spmi.2017.07.001
43. Padmanaban B, Sathiyamoorthy S, Ramesh R (2016) Modeling and simulation of gate engineered gate all-around MOSFET for biomolecule detection. Indian J Sci Technol 9(43):1–6
44. Gautam R, Saxena M, Gupta RS, Gupta M (2013) Gate-all-around nanowire MOSFET with catalytic metal gate for gas sensing applications. IEEE Trans Nanotechnol 12(6):939–944
45. Shahhoseini A (2015) Gas sensor based on silicon nanotube MOSFET. In: Proceedings of 5th International Congress on Nanoscience & Nanotechnology (ICNN2014) 22–24 October 2014, Tehran, Iran, no. October 2014, pp 3–5
46. Kumar N, Raman A (2019) Performance assessment of the charge-plasma-based cylindrical GAA vertical nanowire TFET with impact of interface trap charges. IEEE Trans Electron Devices 66(10): 4453–4460
47. Goel A, Rewari S, Verma S, Gupta RS (2019) Temperature-dependent gate-induced drain leakages assessment of dual-metal nanowire field-effect transistor - Analytical model. IEEE Trans Electron Devices 66(5):2437–2445. https://doi.org/10.1109/TED.2019.2898444
48. Goel A, Rewari S, Verma S, Gupta RS (2021) Modeling of shallow extension engineered dual metal surrounding gate (SEE-DM-SG) MOSFET gate-induced drain leakage (GIDL). Indian J Phys 95(2): 299–308. https://doi.org/10.1007/s12648-020-01704-8
49. Goel A, Rewari S, Verma S, Gupta RS (2021) Novel dual-metal junctionless nanotube field-effect transistors for improved analog and low-noise applications. J Electron Mater 50(1):108–119. https://doi.org/10.1007/s11664-020-08541-9
50. Goel A, Rewari S, Verma S, Gupta RS (2020) High-K spacer dual-metal gate stack underlap junctionless gate all around (HK-DMGS-JGAA) MOSFET for high frequency applications. Microsyst Technol 26(5):1697–1705. https://doi.org/10.1007/s00542-019-04715-6
51. Goel A, Rewari S, Verma S, Gupta RS (2020) Physics-based analytic modeling and simulation of gate-induced drain leakage and

linearity assessment in dual-metal junctionless accumulation nanotube FET (DM-JAM-TFET). Appl Phys A Mater Sci Process 126(5). https://doi.org/10.1007/s00339-020-03520-7
52. Goel A, Rewari S, Verma S, Gupta RS (2018) Dielectric modulated triple metal gate all around MOSFET (TMGAA) for DNA biomolecule detection. Proc. Int. Conf. (2018) IEEE Electron Device Kolkata Conf. EDKCON 2018, vol 1, pp 337–340. https://doi.org/10.1109/EDKCON.2018.8770406
53. Das R, Chanda M, Sarkar CK (2018) Analytical modeling of charge plasma-based optimized nanogap embedded surrounding gate MOSFET for label-free biosensing. IEEE Trans Electron Devices 65(12):5487–5493
54. Hueting RJE, Rajasekharan B, Salm C, Schmitz J (2008) The charge plasma P-N diode. IEEE Electron Device Lett 29(12): 1367–1369
55. Shan C, Wang Y, Bao MT (2016) A charge-plasma-based transistor with induced graded channel for enhanced analog performance. IEEE Trans Electron Devices 63(6):2275–2281
56. Raman A, Kakkar D, Bansal M, Kumar N (2019) Design and performance analysis of GAA Schottky barrier-gate stack-dopingless nanowire FET for phosphine gas detection. Appl Phys A Mater Sci Process 125(11):1–11, [Online]. https://doi.org/10.1007/s00339-019-3066-y
57. Larson JM, Snyder JP (2006) Overview and status of metal S/D Schottky-barrier MOSFET technology. IEEE Trans Electron Devices 53(5):1048–1058
58. Ouarghi M, Dibi Z, Hedjazi N (2018) Impact of triple-material gate and highly doped source/drain extensions on sensitivity of DNA biosensors. J Comput Electron 17(4):1797–1806. https://doi.org/10.1007/s10825-018-1228-3
59. Das A, Kanaujia BK, Nath V, Rewari S, Gupta RS (2020) Impact of reverse gate oxide stacking on gate all around tunnel FET for high frequency analog and RF applications. In: IEEE 17th India Council International Conference (INDICON), pp 1–6

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

2021 Fifth World Conference on Smart Trends in Systems Security and Sustainability (WorldS4) | 978-1-6654-0096-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/WorldS451998.2021.9513992

RPNDroid: Android Malware Detection using Ranked Permissions and Network Traffic

Madan Upadhayay*, Ashutosh Sharma†, Gourav Garg‡ and Anshul Arora§
Discipline of Mathematics and Computing, Department of Applied Mathematics
Delhi Technological University, Delhi, India
*madanraj120999@gmail.com, †ashutosh8110@gmail.com, ‡gauravgarg4000@gmail.com, §anshul15arora@gmail.com

***Abstract*—The number of malware attacks on the Android platform has escalated over the past few years. The seriousness of these attacks can be depicted from the fact that around 25 million Android smartphones were infected with malware within the first half of 2019. Keeping these threats in mind, we aim to develop a hybrid Android malware detector based on ranked permissions and network traffic features. In this work, first, we find the permissions that are frequently present in normal and malicious apps and rank the permissions based upon their frequency in normal and malware dataset. Additionally, we applied different support thresholds to remove the unnecessary and redundant permissions from the rankings. Further, we merge the ranked permissions with the network traffic features to form a hybrid vector for all the apps in the dataset. Finally, we propose a novel algorithm that applies machine learning algorithms on the hybrid vectors consisting of permissions and traffic features to detect Android malware. The experimental results demonstrate that by using the hybrid approach, we could achieve 95.96% detection accuracy with the proposed algorithm.**

***Index Terms*—Android Malware, Permissions, Traffic Features, Hybrid Detection, Malware Detection**

I. Introduction

Smartphones, similar to a computer, are built on a mobile operating system with advanced computing capability and connectivity, and hence their sales have outperformed the desktops' sales [1]. Amongst all mobile OS, Android has been the best market seller due to its open architecture, and a large number of apps in the Play Store. There are 2.46 million apps on Google Play Store and 1.96 million apps on the iOS App Store as of the second quarter of 2019 [2]. This clearly shows the popularity of Android. The presence of the application stores such as Google Play and other third-party app markets makes it easier for the users to download apps. However, not all apps from those markets are *safe*. Some apps, known as malware, may disrupt the normal operation of the smartphone by damaging the system, gather and leak sensitive information, or can cause a financial loss by sending SMS in the background without users' knowledge. The malicious apps can get into the smartphones by one of the three ways: (i) Repackaging in Google Play Store, (ii) Update Attacks, and (iii) Drive-by-Downloads. According to a report [3], 5.3 million malicious mobile installation packages were detected in 2018. Moreover, there were 66.4 million attacks using malicious mobile apps in 2017 which then almost doubled to 116.5 million in 2018 [3]. A malware named xHelper [4] infected around 45000 Android smartphones in the year 2019.

A. Motivation

Keeping the threats posed by the Android malware to the user community in mind, in this work, we aim to propose a hybrid detection mechanism that combines both static permissions and dynamic network traffic features for malware detection. Android provides permissions access as one of the security control mechanisms that provides the restrictions on the certain operations an app can perform. The app developers need to explicitly declare the permissions required by their corresponding app. The list of required permissions is prompted to the user whenever a new app is installed on the smartphone. Generally, users ignore the permissions and grant their access to the application. Malware developers take advantage of these weaknesses and integrate dangerous permissions to generate malicious samples. Hence, out of all static features, we have used permissions in this work for malware detection.

However, many stealthier malware evade static detection because of the presence of update attacks capability. Such samples, initially, do not perform any malicious activity. But they download their malicious component at update time. Because static detection does not aim to analyze the run time or update time behavior of apps, hence, we also analyze the dynamic network traffic features of applications with an aim to detect such stealthier malware. However, not all samples generate network traffic, for instance, a malware might only send SMS in the background without generating any network traffic. Hence, network traffic cannot detect such samples. Therefore, both static permissions and dynamic traffic features have their own limitations. Hence, in this work, we aim to combine both of them to effectively detect Android malware.

Several works exist in the literature that have used permissions for Android malware detection such as [5], [6], [7], [8], [9], [10], etc. Similarly, there exist other related works that have used network traffic features to detect Android malware such as [23], [24], [25], [26], etc. However, both types of detection mechanisms have their own limitations, as discussed above. There exist few works that have used the hybrid combination of permissions and traffic features for detection such as [29], [30], [34], [35], etc. However, none of these hybrid-based works have ranked the permissions for malware detection. They have used all the permissions along with traffic features for detection. We believe that malware developers integrate unnecessary permissions in their malicious samples

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 09:27:14 UTC from IEEE Xplore. Restrictions apply.

to evade detection. Hence, rather than using all the permissions in the detection approach, we should eliminate the unnecessary permissions to improve the detection accuracy.

Hence, in this work, contrary to other hybrid-based works, we aim to rank the permissions to find the minimal set of permissions that yield better detection accuracy. Rather than using all the permissions for detection, we aim to use only the top-ranked distinguishing permissions by ranking them in order of their frequency difference in malware and normal dataset. We obtain two separate permission rankings: one representing the permissions that are significantly present in malware apps; and second representing the permissions that are significantly present in the normal dataset. Moreover, we also take the top ranked traffic features from one of our previous works [26]. We further merge the top ranked permissions with ranked traffic features to construct a hybrid vector. To the best of our knowledge, no other work has used the ranked permissions, and ranked traffic features for hybrid Android malware detection.

B. Contributions

The main contributions of this work are highlighted below:

1) We ranked the permissions based upon their frequency in normal and malware dataset.
2) We obtained two different permission rankings: one for frequently occurring permissions in malware and other for frequently occurring permissions in normal apps.
3) We applied several support thresholds to eliminate the lower-ranked permissions.
4) We further merged the top ranked permissions with the ranked traffic features to create a hybrid feature vector.
5) We applied machine learning algorithms on the hybrid feature vectors to detect Android malware.

C. Organization

The remainder of the paper is structured as follows. Section II reviews the existing works in the literature for Android malware detection. The detailed methodology is explained in Section III. We present the results obtained from the proposed *RPNDroid* model in Section IV, and we conclude with future work directions in Section V.

II. Related Work

In this section, we discuss the solutions proposed for Android malware detection. Broadly, such solutions can be categorized into three types: Static Solutions, Dynamic Solutions, and Hybrid Solutions. The following subsections discuss these detection mechanisms.

A. Static Solutions

In static solutions, the researchers aim to detect Android malware without executing the apps on the smartphone. The researchers aim to analyze the static components of an application such as permissions, intents, Java Code, API calls, and other manifest file components.

1) Permissions Based Detection: Earlier works such as [11], [12] analyzed the presence of dangerous permission patterns in normal apps. The authors in [7] observed the permissions and their patterns present in malicious apps, without focusing on detection. Wang et al. [6] ranked the permissions according to the Information Gain, and T-Test, and applied machine learning algorithms. The authors in [9] evaluated the permissions' score based upon its frequency in the malware dataset and used that score for detection. The authors in [10] also performed permissions' ranking based upon their frequency in malware and normal dataset for detection. The authors in [8] used permission pairs for detection.

2) Manifest File Based Detection: Some of the works have used other components of the manifest file such as intents, hardware components, network addresses and strings for malware detection. Sanz et al. in their works [5] and [13] applied machine learning on manifest file components including hardware components, intents, and permissions. Arp et al. [14] proposed a model based on all of the manifest components for malware detection. Kim et al. [16] constructed a static feature vector for all the apps consisting of manifest components and API calls and further used those vectors for malware detection. The authors in [15] proposed a model named *DroidMat* that applied K-means clustering on permissions, intents, APIs, and hardware components. The works proposed in [17] and [36] have used the combination of intents with permissions to detect malware.

3) API Calls Based Detection: Elish et al. [18] studied sensitive APIs to observe user-triggered dependence for malware detection. The authors in [19] established the dependency graphs of the API calls generated from the dataset and used similarity metrics to detect malicious activity. However, sometimes, if the APIs used for analysis or detection are not connected to any of the manifest component, then these APIs may act as noise in the detection process. Therefore, we did not aim to study APIs or Java code to detect Android malware in this work.

B. Dynamic Solutions

Malicious components, in some of the malware, may get downloaded at run time or update time. Since static solutions do not aim to analyze the run time behavior of apps, hence, static solutions may not detect such malware. Therefore, researchers focused on dynamic solutions that aim to execute the apps and analyze their run time behavior. Dynamic solutions are categorized into two types: OS-based (system calls based) detection and Network Traffic based detection.

1) OS Based Detection: Reina et al. [20] studied system calls of malicious apps and further mapped those calls to run time JNI, Java, and native code execution. Shabtai et al. [21] analyzed a mix of dynamic features like the amount of CPU utilization, processes running in the background, and packets transmitted over the network, to detect Android malware. Afonso et al. [22] analyzed dynamic API calls and system calls for malware detection. It has been observed that system calls as a feature gives relatively lower detection accuracy for

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 09:27:14 UTC from IEEE Xplore. Restrictions apply.

Android malware. Hence, we have not used system calls in this work.

2) Network Traffic Based Detection: Shabtai et al. [23] studied network traffic behavior of benign dataset. The authors further applied machine learning algorithms for detecting deviations from the normal patterns in the malicious traffic. The authors in [24] observed the patterns in the HTTP headers generated from network traffic for malware detection. Arora et al., in their two previous works, [25] and [26] analyzed network traffic of malicious and normal apps and further used rule based classifier and Naive Bayes model respectively to detect malware. However, all these traffic based mechanisms can be useful only if the malware samples are controlled by some remote server. Hence, we have merged the network traffic with permissions to detect Android malware.

C. Hybrid Solutions

Some of the detection mechanisms construct a hybrid vector by combining the static and dynamic features, and hence, propose a hybrid model. Grace et al. [28] proposed *Riskranker*, a detection model that observed static permissions and dynamic features like run-time Dalvik code loading for malware detection. Saracino et al. [27] analyzed various static as well as dynamic features including user activities, system calls, permissions, etc., for malware detection. The authors in [29] and [30] combined network traffic features with static permissions in their works for malware detection. However, none of the above discussed works have aimed to rank both permissions and traffic features for detection. In this work, we aim to eliminate the unnecessary permissions first before merging them with traffic features for effective malware detection.

III. Proposed Methodology

Now, we present the proposed model in detail which is divided in several phases. We explain all the phases in the upcoming subsections.

A. Permissions Ranking

First, we aim to rank the permissions so that we can use the top-ranked distinguishing permissions in the detection phase along with traffic features. The permissions required by any application are defined in its manifest file. We used *APKTOOL* to break the apks and further extract their manifest files. Next, we wrote scripts in Python to extract the permissions required by each application from its corresponding manifest file.

We represented the permissions extracted from the normal dataset in the normal permission matrix represented by *B*, and the malicious permissions in the malware permission matrix, say *M*. The rows and the columns in the matrices *B* and *M* represent the applications and the permissions respectively. The entries in the matrices are either 0 or 1. The entry at any index [i,j] is 1 if the ith app contains the jth permission, else the entry is 0. Because the number of apps in the normal and malware dataset may vary, we need to make balanced matrices. Hence, we calculate the relative frequency of every permission in both the matrices. To calculate the relative frequency of any permission in the normal matrix, we divide the frequency of that permission in the normal apps by the total number of normal apps. On similar lines, we divide the frequency of every permission in the malware dataset by the total number of malware apps to find the relative frequency of every permission in the malware matrix. Let the relative frequency of any j^{th} permission is denoted by P_j and Q_j respectively in benign and malware dataset.

We need to find the permissions that are significantly present in malware apps but not in normal ones, and vice-versa. We use the term *dangerous permissions* to denote the permissions that are significantly present in malware apps. Similarly, we define the term *clean permissions* that are significantly present in normal apps. To find the *dangerous* and *clean* permissions, we find a score for every unique permission using the following equation.

$$R(P_j) = Q_j - P_j \tag{1}$$

For every permission, the value of $R(P_j)$ will lie in the closed interval [-1 , 1]. If the value of $R(P_j)$ is 1, it means the permission is present only in the malware dataset and hence it is the most *dangerous* permission. The value of $R(P_j) = -1$ signifies the most *clean* permission. Further, we rank all the permissions based upon their scores obtained. We create two ranked lists, one in ascending order of their scores represented by A_{List} and the other one in descending order represented by D_{List}. The top permissions in the A_{List} are the ones that are significantly present in the normal dataset. And the top permissions in the D_{List} are the ones that are significantly present in the malware dataset.

B. Permissions Threshold

There are many permissions in the Android system and some of them may not be useful in detecting malicious apps. Hence, after ranking the permissions, we aim to eliminate the lower-ranked permissions from both A_{List} and D_{List}. We prepare different threshold limits ranging from 5% to 50% in the interval of 5. Any limit of *k*% in the interval [5,50] denotes that we consider those permissions for our experiments that are present in at least *k*% of both normal and malware dataset and eliminate the rest of the permissions. We evaluate our model on each of the threshold values in the interval [5, 50].

C. Hybrid Feature vectors

Next, we merge the top-ranked permissions with network traffic features to construct a hybrid feature vector. In one of our previous works [26], we ranked 22 traffic features and observed that the top nine features give better accuracy as compared to all the features for detection. Hence, in this work, we have used the same top nine features obtained from our previous work. Table I summarizes the nine features used in our work. To extract the traffic features from the apps, we used Samsung Galaxy smartphones and installed the "tPacketCapture" application on them to capture the network traffic. We infected the smartphones with each malware for 96 hours to get the malware traffic. Normal traffic was captured

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 09:27:14 UTC from IEEE Xplore. Restrictions apply.

TABLE I: Traffic Features used in Experiments.

Average Packet Size (F1)	Ratio of Incoming to Outgoing Bytes (F2)	Packets Sent per Second (F3)
Minimum Time Interval between Packets Sent (F4)	Minimum Time Interval between Packets Received (F5)	Packets Received per Second (F6)
Packets Sent per Flow (F7)	Packets Received per Flow (F8)	Maximum Packet Size (F9)

from the normal web browsing on the Android smartphone. We wrote shell scripts to extract the traffic features from the captured network traffic files.

To construct a hybrid feature vector for any app, we merge the permissions of the app and the numerical values of nine traffic features corresponding to that app. Note that we have ranked the permissions and further applied support threshold values to eliminate the lower-ranked permissions. Hence, if we assume that there are *N* number of total permissions in any app, and we apply *k%* of support threshold on A_{List} and D_{List}, then a subset of permissions will be selected from that app that are above the support threshold. Let that subset of permissions be of size *S*, where *S* is less than or equal to *N*. We extract nine traffic features for that app. Therefore, when we merge the selected permissions and traffic features to construct a hybrid vector for that particular app, the vector will be of the size *"S+9"*. The value of *S* will vary when we change the support threshold. Moreover, because different apps contain the different number of permissions, the hybrid vector will be of different sizes for the apps.

D. Testing Phase

In the previous phase, say, we select the top-ranked *k%* permissions, and merge them with traffic features to construct a hybrid vector. Let the hybrid feature vector of any application be represented as a set $V = \langle P_1, P_2, P_3, ..P_k, F_1 = \alpha 1, F_2 = \alpha 2, ...F_9 = \alpha 9\rangle$. Here, $\langle P_1, P_2, P_3, P_4, ..P_k\rangle$ represent the ranked permissions that are above the *k%* threshold. And rest are the numerical values of the nine traffic features. Similarly, we obtain the hybrid vectors for all the apps in the malware and normal dataset. Next, we apply several machine learning classifiers like Naive Bayes, Random Forest, and SVM to train on the hybrid vectors of normal and malware dataset.

Now, to test any app, say A_T, we extract all of its permissions and values of nine traffic features and construct its hybrid vector. It might happen that some permissions of A_T might not be present in the trained hybrid vectors because such permissions might not exceed the *k%* support threshold. Hence, we ignore all such permissions and do not consider them in detection. Note that we use this threshold only on permissions and not on traffic features because there are many unnecessary permissions that can hinder the detection accuracy. We extract all the nine traffic features for all the testing apps. We apply machine learning algorithms on the extracted hybrid vector of A_T to determine whether A_T is malicious or benign.

IV. Results and Discussion

Now, we present the results obtained from the *RPNDroid* model. We used 2500 malware samples from various sources such as Genome [31], Drebin [32], and Koodous [33]. We downloaded 2400 normal apps from the official Google Play Store. Further, to make sure the downloaded apps from the Play Store are benign and not malicious, we tested these apps on the VirusTotal, that hosts various anti-malware apps such as Kaspersky, Norton, McAfee, Security 360, etc.

TABLE II: Top 15 Permissions in Ascending and Descending Order.

Ascending Order (Top Clean Permissions)		**Descending Order (Top Dangerous Permissions)**	
Permission	**Score**	**Permission**	**Score**
WAKE LOCK	-0.313938	READ PHONE STATE	0.488382
GET ACCOUNTS	-0.253409	READ SMS	0.456050
READ EXTERNAL STORAGE	-0.228872	SEND SMS	0.433686
USE CREDENTIALS	-0.170744	WRITE SMS	0.395856
ACCESS NETWORK STATE	-0.158858	RECEIVE SMS	0.352381
CAMERA	-0.121557	RECEIVE BOOT COMPLETED	0.269169
VIBRATE	-0.087729	CHANGE WIFI STATE	0.228281
AUTHENTICATE ACCOUNTS	-0.053013	READ CONTACTS	0.203704
RECORD AUDIO	-0.045398	CALL PHONE	0.192382
MANAGE ACCOUNTS	-0.044722	RESTART PACKAGES	0.188282
READ SYNC SETTINGS	-0.042529	WRITE APN SETTINGS	0.185852
WRITE SYNC SETTINGS	-0.040887	WRITE CONTACTS	0.178911
BLUETOOTH	-0.038583	GET TASKS	0.164140
READ APP BADGE	-0.037411	MOUNT UNMOUNT FILESYSTEMS	0.145570
INTERNET	-0.024915	INSTALL PACKAGES	0.143677

Before we highlight the detection results, first we summarize, in Table II, the top 15 permissions obtained from the relative frequency difference in ascending as well as descending order. For every permission, the table also summarizes the score obtained for that permission, i.e., the difference in relative frequencies for that permission in malware and normal dataset. This list of top 15 permissions is obtained at the 0% support threshold. As can be seen from Table II, no permission gets the score of +1 or -1 which means there is no permission that is present only in one type of the dataset.

Now, we discuss the detection accuracy obtained from the proposed approach. We review the results in three subsections: a). Static Results, i.e., if we use ranked permissions alone for detection, b). Dynamic Results, i.e., we use network traffic features for detection, and c). Hybrid Results, i.e., the combination of ranked permissions and traffic features for detection.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 09:27:14 UTC from IEEE Xplore. Restrictions apply.

TABLE III: Static Detection Results with Different Classifiers at Different Thresholds (Permissions Ranked in Ascending Order).

Threshold (in%)	Machine Learning Algorithms Accuracy (%)		
	Naive Bayes	Random Forest	SVM
0	83.17	91.34	89.38
5	84.07	88.75	88.68
10	82.96	83.44	83.31
15	74.37	80.86	80.72
20	74.37	81.14	81.21
25	74.37	81.14	81.21
30	74.37	81.14	81.21
35	74.37	75.83	75.83
40	58.44	61.38	61.59
45	58.44	61.38	61.59
50	58.44	58.37	58.51

TABLE IV: Static Detection Results with Different Classifiers at Different Thresholds (Permissions Ranked in Descending Order).

Threshold (in%)	Machine Learning Algorithms Accuracy (%)		
	Naive Bayes	Random Forest	SVM
0	86.52	91.41	91.96
5	87.08	91.96	90.50
10	83.65	84.84	84.91
15	79.32	81.56	81.63
20	74.86	74.86	74.93
25	74.86	74.86	74.93
30	74.86	74.86	74.93
35	74.86	74.86	74.93
40	58.37	61.38	61.52
45	58.37	61.38	61.52
50	58.37	58.51	58.51

A. Static Detection Results

First, we review the results if we use ranked permissions alone for detection. We apply different machine learning classifiers at different support thresholds of permissions. Since we have two separate ranked permissions lists, in ascending and descending order, we discuss the results obtained from each of the ranked lists in the following separate subsections.

1) Results from Ranked Permissions in Ascending Order: First, we discuss the results obtained from the top clean permissions that are ranked higher in the ascending order of the permissions' score. Table III summarizes the detection results. As Table III shows, Naïve Bayes gives the highest accuracy of 84.07% at 5% support threshold. Random Forest and SVM give relatively better accuracy than Naïve Bayes. We achieve the highest accuracy of 91.34% with the Random Forest classifier at 0% threshold.

2) Results from Ranked Permissions in Descending Order : Now, we discuss the results obtained from the top dangerous permissions that are ranked higher in the descending order of the permissions' score, summarized in Table IV. As Table IV highlights, Naïve Bayes gives the highest accuracy of 87.08% at 5% threshold. Similar to the case of ascending order results, in this case as well, Random Forest and SVM give relatively better accuracy than Naïve Bayes. We achieve the highest accuracy of 91.96% with Random Forest and SVM at 5% and 0% threshold respectively.

We observe that stealthier samples like *BaseBridge* and *AnserverBot* evade static detection because of presence of update attacks capability. Hence, we focus on dynamic results in the next subsection.

TABLE V: Dynamic Detection Results with Nine Traffic Features.

	Machine Learning Algorithms		
	Naive Bayes	Random Forest	SVM
Detection Accuracy (%)	94.27	94.76	94.76

TABLE VI: Hybrid Detection Results with Different Classifiers at Different Thresholds (Permissions Ranked in Ascending Order).

Threshold (in%)	Machine Learning Algorithms Accuracy (%)		
	Naive Bayes	Random Forest	SVM
0	86.27	95.34	94.48
5	87.27	92.64	92.88
10	86.76	88.44	88.44
15	79.67	84.96	84.96
20	79.67	85.24	84.41
25	79.67	85.24	84.41
30	79.67	85.24	84.41
35	79.67	79.33	79.33
40	61.74	64.48	65.79
45	61.74	64.48	65.79
50	61.74	64.48	64.48

B. Dynamic Detection Results

Next, we review the detection results if we use only the nine traffic features for malware detection. Table V summarizes the dynamic detection results. In our previous work [26], the proposed model gave us the nine best set of ranked features. Hence, in this work, we have not ranked the traffic features again and have used the same set of features. With the nine features (mentioned in Table I), we get the detection accuracy of above 94% with all the three classifiers (Table V). We can conclude that dynamic traffic features give better results than static permissions.

However, some samples do not generate any network traffic, and hence cannot be detected by a traffic-based approach. Hence, we aim to integrate the ranked permissions with the ranked features to further improve the detection accuracy.

C. Hybrid Detection Results

Finally, we evaluate the results with hybrid feature vectors consisting of both permissions and traffic features. Tables VI and VII summarize the results when permissions are ranked in ascending and descending order respectively.

We observe from Table VI that when permissions are ranked in ascending order and merged with nine traffic features, then

TABLE VII: Hybrid Detection Results with Different Classifiers at Different Thresholds (Permissions Ranked in Descending Order).

Threshold (in%)	Machine Learning Algorithms Accuracy (%)		
	Naive Bayes	Random Forest	SVM
0	88.62	94.81	95.96
5	90.81	95.96	94.50
10	86.75	87.74	88.90
15	81.52	83.66	83.83
20	76.56	76.56	76.63
25	75.56	75.56	75.73
30	75.56	75.56	75.73
35	75.56	75.56	75.73
40	62.47	64.78	65.72
45	61.57	62.48	62.52
50	61.57	61.77	61.77

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 09:27:14 UTC from IEEE Xplore. Restrictions apply.

we get the best detection accuracy of 95.34% with Random Forest classifier at 0% threshold. Similarly, from Table VII, with permissions ranked in descending order, Random Forest and SVM give the accuracy of 95.96% at 5% and 0% threshold respectively. Hence, we can say that the hybrid approach gives us the accuracy of 95.96%, better than achieved from static and dynamic features.

V. Conclusions

In this work, we proposed a hybrid Android malware detection technique based on ranked permissions and traffic features. We considered the frequency of the permissions present in the normal and malware dataset as the basis for ranking the permissions. We created two ranked lists: one list of significant permissions present in normal apps; and the other of significant permissions present in malware apps. Further, we also applied the support threshold concept to eliminate the lower-ranked permissions. We took prioritized nine traffic features from one of our previous works and merged the traffic features with ranked permissions to create a hybrid vector for all the apps. Finally, we applied several machine learning algorithms on the hybrid vectors to detect Android malware. The results demonstrated that the hybrid approach of merging the permissions with traffic features is better than using permissions or traffic features alone. In our future work, we will aim to include other manifest file components as well such as intents and hardware components in our analysis.

Acknowledgment

The authors would like to thank the researchers of Genome, Drebin and Koodous projects for providing us the malware dataset for our experiments.

References

[1] Desktop vs Mobile vs Tablet Market Share Worldwide (Dec 2018 - Dec 2019), Available Online. "https://gs.statcounter.com/platform-market-share/desktop-mobile-tablet".
[2] Number of apps available in leading app stores as of 3rd quarter 2019, Available Online. "https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/".
[3] Mobile malware evolution 2018, Available Online. "https://securelist.com/mobile-malware-evolution-2018/89689/".
[4] 45,000 Android devices infected by new unremovable xHelper malware, Available Online. "https://thenextweb.com/security/2019/10/30/45000-android-devices-infected-by-new-unremovable-xhelper-malware/".
[5] B. Sanz et al., Puma: Permission Usage to Detect Malware in Android, International Joint Conference CISIS'12-ICEUTE 12-SOCO 12 Special Sessions, Springer Berlin Heidelberg, 2013.
[6] W. Wang et al., Exploring Permission-Induced Risk in Android Applications for Malicious Application Detection, IEEE Transactions on Information Forensics and Security, vol. 9 , pp. 1869-1882, 2014.
[7] V. Moonsamy, J. Rong, and S. Liu, Mining Permission Patterns for Contrasting Clean and Malicious Android Applications, Future Generation Computer Systems, vol. 36, pp. 122-132, 2014.
[8] A.Arora, S.K. Peddoju, and Mauro Conti, PermPair: Android Malware Detection using Permission Pairs, IEEE Transactions on Information Forensics and Security, vol. 15, pp. 1968-1982, 2020.
[9] K. Talha, D. Alper, and C. Aydin, APK Auditor: Permission-based Android Malware Detection System, Digital Investigation, vol. 13, pp. 1-14, 2015.
[10] J. Li et al., Significant Permission Identification for Machine-Learning-Based Android Malware Detection, IEEE Transactions on Industrial Informatics, vol. 14, pp. 3216-3225, 2018.
[11] M. C. Grace, W. Zhou, X. Jiang, and A. Sadeghi, Unsafe exposure analysis of mobile in-app advertisements, 5th ACM WiSec, 2012.
[12] W. Enck, M. Ongtang, and P. McDaniel, On Lightweight Mobile Phone Application Certification, 16th ACM CCS, 2009.
[13] B. Sanz et al., MAMA: manifest analysis for malware detection in android, Cybernetics and Systems, vol. 44, pp. 469-488, 2013.
[14] D. Arp, M. Spreitzenbarth, M. Hubner, H. Gascon, and K. Rieck, DREBIN: Effective and Explainable Detection of Android Malware in Your Pocket, NDSS, 2014.
[15] D. J. Wu at al., Droidmat: Android malware detection through manifest and api calls tracing, Seventh Asia Joint Conference on Information Security (Asia JCIS), 2012.
[16] T. Kim, B. Kang, M. Rho, S. Sezer, and E. Im, A Multimodal Deep Learning Method for Android Malware Detection Using Various Features, IEEE Transactions on Information Forensics and Security, vol. 14, 2019.
[17] F. Idrees, M. Rajarajan, M. Conti, T. Chen, and Y. Rahulamathavan, PIndroid: A novel Android malware detection system using ensemble learning methods, Computers & Security, vol. 68, pp. 36-46, 2017.
[18] K. Elish et al., Profiling user-trigger dependence for Android malware detection, Computers & Security, vol. 49, pp. 255-273, 2015.
[19] M. Zhang, Y. Duan, H. Yin, and Z. Zhao, Semantics-Aware Android Malware Classification Using Weighted Contextual API Dependency Graphs, ACM CCS, 2014.
[20] A. Reina, A. Fattori, and L. Cavallaro, A System Call-Centric Analysis and Stimulation Technique to Automatically Reconstruct Android Malware Behaviors, 6th European Workshop on System Security, 2013.
[21] A. Shabtai, U. Kanonov, Y. Elovici, C. Glezer, and Y. Weiss, Andromaly: a behavioral malware detection framework for android devices, Journal of Intelligent Information Systems, vol. 38, pp. 161-190, 2011.
[22] V.M. Afonso et al., Identifying Android malware using dynamically obtained features,Journal of Computer Virology and Hacking Techniques, vol. 11, pp.9-17, 2015.
[23] A. Shabtai et al., Mobile malware detection through analysis of deviations in application network behavior, Computers & Security, vol. 43, pp. 1-18, 2014.
[24] S. Wang, et al., Detecting Android Malware Leveraging Text Semantics of Network Flows, IEEE Transactions On Information Forensics And Security, vol. 13, pp. 1096-1109, 2018.
[25] A. Arora, S. Garg, and S. K. Peddoju, Malware detection using network traffic analysis in android based mobile devices, 8th IEEE NGMAST, 2014.
[26] A. Arora, and S. Peddoju, Minimizing Network Traffic Features for Android Mobile Malware Detection, 18th ACM ICDCN, 2017.
[27] A. Saracino, D. Sgandurra, G. Dini, and F. Martinelli, MADAM: Effective and Effcient Behavior-based Android Malware Detection and Prevention, IEEE Transactions on Dependable and Secure Computing, vol. 15, pp. 83-97, 2018.
[28] M. Grace, Y. Zhou, Q. Zhang, S. Zou, and X. Jiang, Riskranker: scalable and accurate zero-day android malware detection, 10th ACM Mobisys, 2012.
[29] A. Arora, and S. Peddoju, NTPDroid: A Hybrid Android Malware Detector Using Network Traffic and System Permissions, 17th IEEE TrustCom, 2018.
[30] A. Arora, S. Peddoju, V. Chauhan, and A. Chaudhary, Hybrid Android Malware Detection by Combining Supervised and Unsupervised Learning, 24th ACM MobiCom, 2018.
[31] Y. Zhou, and X. Jiang, Dissecting Android Malware: Characterization and Evolution, IEEE Symposium on Security and Privacy, 2012.
[32] D. Arp, M. Spreitzenbarth, M. Hubner, H. Gascon, and K. Rieck, DREBIN: Effective and Explainable Detection of Android Malware in Your Pocket. NDSS, 2014.
[33] Koodous Malware Dataset, "www.koodous.com".
[34] Y. Shyong, T. Jeng, and Y. Chen, Combining Static Permissions and Dynamic Packet Analysis to Improve Android Malware Detection, 2nd International Conference on Computer Communication and the Internet (ICCCI), 2020.
[35] S. Kandukuru, and R. M. Sharma, Android malicious application detection using permission vector and network traffic analysis, 2nd International Conference for Convergence in Technology (I2CT), 2017.
[36] K. Khariwal, J. Singh and A. Arora, IPDroid: Android Malware Detection using Intents and Permissions, Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), UK, 2020.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 09:27:14 UTC from IEEE Xplore. Restrictions apply.

Rendiconti Lincei. Scienze Fisiche e Naturali
https://doi.org/10.1007/s12210-021-01017-w

RESEARCH PAPER

Seasonal variability and kinetics of phosphate removal in a *Phragmites*-based engineered wetland

Harsh Pipil[1] · A. K. Haritash[1] · Krishna R. Reddy[2]

Received: 27 May 2021 / Accepted: 31 July 2021
© Accademia Nazionale dei Lincei 2021

Abstract
Phosphate is a critical pollutant leading to eutrophication in water bodies and significant adverse effect on water quality around the globe. The present study assessed the role of *Phragmites*-based constructed wetland (CW) towards the removal of phosphate from wastewater under natural environment and to assess the influence of meteorological parameters and seasonal changes. The *Phragmites*-based wetland removed phosphate with average efficiency of 80% at HRT of 24 h, and the removal efficiency increased with ambient temperature and amendment of ammonium in wastewater. Rhizospheric environment with slightly reducing conditions ($Fe^{3+}/Fe^{2+} < 1.0$) and pH less than 8.0 do not favour binding of phosphate with Fe and Ca/Mg, respectively. The study concluded that *Phragmites* spp. plays the dominant role in phosphate removal (based on kinetics of removal) up to 20 mg/l-PO_4^{3-} in wastewater. While the treatment is observed without input of chemicals or energy, it is a sustainable eco-friendly option for decentralised wastewater treatment throughout the year.

Keywords *Phragmites* · Constructed wetland · Eutrophication · Phosphate · Kinetics

1 Introduction

India is the second most populous country in the world that produces large volumes of wastewater, and improper discharge practices have resulted in pollution in most of the water bodies around the country. Among the contaminants/nutrients present in wastewater/runoff, phosphate is a crucial factor which regulates algal-blooms, and the level of cultural eutrophication in freshwater ecosystems, particularly in agriculture-dominated developing countries like India (Shan et al. 2020, 2021). Whereas the conventional treatment of wastewater is less efficient and energy intensive, the advanced processes that use iron, aluminum, and calcium-based materials for phosphate removal are chemical intensive and produce secondary sludge. Therefore, there is a high demand for a treatment technology which has low operating and maintenance cost, produces no secondary sludge, and is environment-friendly (Shaltout et al. 2014; Binder et al. 2017; Haritash et al. 2015; Talukdar and Pal 2017). The constructed wetlands (CWs) have been reported to remove pollutants such as total suspended solids (TSS), pathogens, metals, colour, nutrients, etc., and help mitigating the environmental effects of eutrophication (Schmitt et al. 2015a, b; Al-Isawi et al. 2017). The wetlands operate on solar energy with no input of chemicals or technical expertise; and the factors that affect efficiency and performance of CW are its area, length to width ratio, depth of wastewater in CW cell, wastewater loading rate, and hydraulic retention time (Pant et al. 2001). The substrate used for plant growth in CW has also been reported to affect the pollutant removal through filtration, but its efficiency gradually reduces with saturation and clogging (Kajla et al. 2021). Apart from filtration, absorption, adsorption, and biochemical precipitation are the major mechanisms for pollutant removal in wetlands. On the other hand, the removal of pollutants by wetland vegetation remains unaffected in tropical and sub-tropical regions since plants remove pollutants through absorption and their activity/evapotranspiration is regulated by climate/

✉ A. K. Haritash
akharitash@dce.ac.in

Harsh Pipil
harsh_phd2k19@dtu.ac.in

Krishna R. Reddy
kreddy@uic.edu

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

[2] Department of Civil, Materials, and Environmental Engineering, University of Illinois, Chicago, IL 60607, USA

Published online: 14 August 2021

temperature of that region (Singh et al. 2015) which remains warm most of the time during a year. Some studies have confirmed effective removal of pollutants/nutrients from wastewater throughout the year, particularly in subtropical regions (Nandakumar et al. 2019).

The novelty of the present study lies in determining the phosphate removal efficiency of a Phragmites-based constructed wetland throughout the four seasons in India. The volume and composition of wastewater changes considerably during different seasons, so, it is important to investigate the kinetics of removal under varying load of nutrients (PO_4^{3-}). The treatment of wastewater through CWs is of significant environmental importance since the biochemical transformation of organic impurities/pollutants, by microbes and vegetation in CW, into harmless/mineralised forms lead to incorporation of nutrients into the food chain, thus, resulting in ecosystem services (Camselle and Gouveia 2019; Singh et al. 2019). Since plant uptake is primarily regulated by weather conditions, it is important to study the effect of ambient temperature on the removal efficiency of CWs. Based on these considerations, the present study aims to exploit the natural tendency of plants (*Phragmites* spp.) towards phosphate removal from high strength wastewater (Metcalf and Eddy 2003) in relation to ambient temperature during four different seasons. It is also aimed to study the kinetics involved in nutrient removal by CWs under different seasons and loads so that supplementary arrangements can be made for wastewater treatment, if necessary.

2 Materials and methods

2.1 Configuration of CW cell

The present study was undertaken in a *Phragmites*-based vertical sub-surface flow constructed wetland located in Delhi Technological University (DTU), Delhi, India. The study was conducted continuously for a period of 7 months to cover the four seasons of autumn, winter, spring, and summer so that the seasonal variations can be investigated. The CW cell was constructed in open, using brick masonry, to carry out the experiments under natural conditions of temperature and sunlight. The CW cell was filled with the mixture of gravel-sand as substrate medium with a depth of 35 cm, length 110 cm, width of 80 cm, and the surface area being 0.88 m^2 (Fig. 1). The approximate volume of the CW cell was 400 L; bed packing material consisted of 39% gravel, 59% sand, and 2% silt, with specific gravity of solids (G_s) of 2.74, void ratio (e) 0.77, and bulk density (ρ) of 1782 kg/m^3. The media bed provided suitable growth environment with easy root penetration, percolation of influent and gas exchange; and the gap-graded size profile of substrate in CW cell ensured more hydraulic conductivity and thus helped in nutrient removal through subsurface flow of wastewater. During April (before summer season), 20 number of *Phragmites* plants were planted in a grid pattern of 4×5 making the plant density of 22.72 plants/m^2 and a considerable time was given to the plants to adapt to its new habitat. Initially, the average height of the plant was 106 cm. At the final stage of the experiment, the number of plants were 180 and the average height was 213 cm.

2.2 Operation of CW cell

After stabilization of *Phragmites* to its new habitat, it was fed with water at the rate of 20 L per day. The di-ammonium phosphate (DAP) and 50 g of urea were also fed into the CW system. In the month of October (autumn season), the system was fed with synthetically prepared wastewater for which analytical grade di-hydrogen orthophosphate (KH_2PO_4) salt was used. The phosphate (PO_4^{3-}-P) concentration in synthetic wastewater was progressively increased

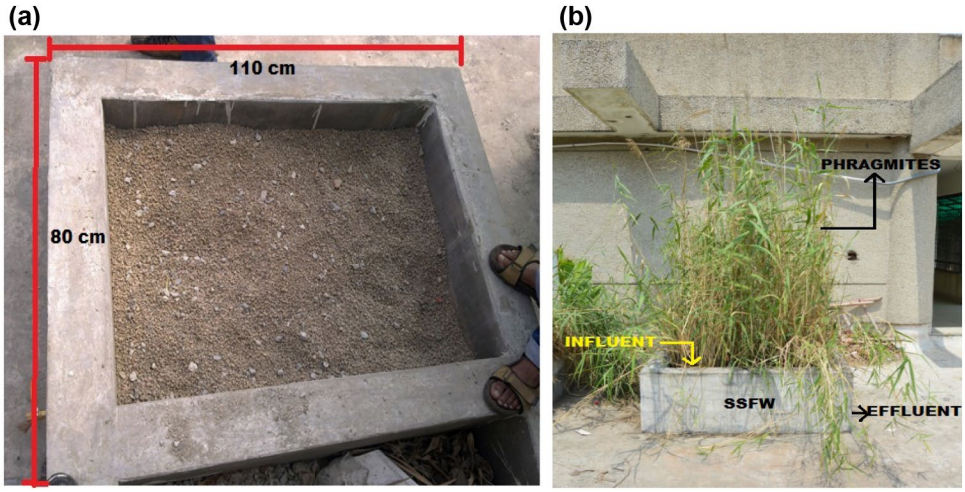


Fig. 1 **a** Prepared substrate bed, **b** and stabilized CW Cell for the study

from 5 mg/l (autumn and winter) to 10 mg/l (spring) followed by 20 mg/l (summer) during the different stages of treatment process. The dose of phosphate was increased considering the rate of evapotranspiration in varying seasons to study removal efficiency of *Phragmites* and its response towards the chemical shocks with respect to phosphate loading. Further increase in phosphate concentration (> 20 mg/l) was avoided since it is very high concentration and is rarely observed in runoff or natural water bodies (Metcalf and Eddy 2003). The 5 mg/l PO_4^{3-}-P in autumn and winter season, 10 mg/l PO_4^{3-}-P and 50 mg/l NH_4^+-N in spring; and 20 mg/l PO_4^{3-}-P and 100 mg/l NH_4^+-N in summer season were added to the CW cell to study the removal of phosphorous and nitrogen. Ammonium ions (NH_4^+) were added to analyse the redox conditions in the CW cell based on nitrification/denitrification process. The synthetic wastewater of varying concentration was added into the system from the inlet valve at top, and the sample was collected from outlet valve at the bottom of the CW cell. The hydraulic retention time (HRT) in the cell was 24 h. The samples collected were analyzed for phosphate as available phosphate (AP), total phosphate (TP), Total Kjeldahl's Nitrogen (TKN) for organic nitrogen, and ammonia to find its presence in the treated wastewater sample. The ratio of ferric to ferrous ion (Fe^{3+}/Fe^{2+}) was also determined to determine the redox state in CW cell. Orion Star A329 model multimeter was used to measure total dissolved solids (TDS) and electrical conductivity (EC), while pH was measured using Hanna Instrument make (Italy) (Model No. HI 9607) pen type pH meter. Phosphate was determined spectrophotometrically using molybdate method, TKN was analysed on automated TKN assembly, and iron was determined by phenanthroline method. The meteorological parameters of ambient temperature, sunshine hours, and rainfall were also recorded from Indian Meteorological Department (IMD), Delhi.

3 Results and discussion

3.1 Meteorological observations and characterization of substrate

The ambient temperature in Delhi, India was noted for a period of 7 months during the study to cover the stretch of four seasons (autumn, winter, spring and summer). The ambient temperature varied from 2 to 42 °C during the study, with average minimum and maximum temperature of 13.6 °C and 26.8 °C, respectively. During the autumn, winter, spring and summer seasons, the minimum and the maximum temperature varied between 9 °C to 36 °C; 2 °C to 28 °C; 7 °C to 34 °C and 17 °C to 42 °C, respectively. The temperature profile represented the sub-tropical climatic conditions for the study area. Average sunshine of 8–12 h was recorded during the study. Precipitation in terms of rainfall (mm) was received in small 18 spells which had negligible influence on the study.

For the substrate, bed packing material contained 39% gravel, 59% sand, and 2% silt (Fig. 2). The gap graded distribution of packing material provided more hydraulic conductivity, easy gaseous exchange, and better root penetration, thus, helping in nutrient removal evenly, as observed in other studies (Vymazal 2005). Further, the substrate profile ensured homogenous nutrient distribution and redox conditions throughout the cell column.

3.2 Removal of available phosphate (AP) and total phosphate (TP)

During the study, the removal efficiency of available phosphate (AP) varied with the seasons and increased with the increasing influent concentration of phosphate. Removal was maximum in summer season, followed by spring, winter, and autumn. The removal efficiency followed the trend of 98.5% > 91.2% > 85.8% > 63.7% for summer, spring, winter, and autumn, respectively (Table 1). It may be attributed to the increase in average sunshine hours and rise in ambient temperature resulting in increased evapotranspiration and the metabolic activity of the plant in summers. Various other mechanisms that are involved in the removal of the phosphates are adsorption by substrate-material and sediments as well as microbial uptake (Kadlec and Knight 1996). Studies have reported that phosphate removal by adsorption over the sediment bed decreases with time due to its saturation (Kajla et al. 2021); while in present study, the efficiency continue to increase (from 63.7% in autumn to 98.5% in summer) till the end of the study indicating that the plants have played a major role in removal of phosphate. It may be concluded that the removal of AP by bed sediments was limited during the initial phase of study as reported in earlier literature by the authors (Haritash et al. 2017). The increase in the sunshine hours led to overall increase in photosynthesis, and the rise in ambient temperature causing more

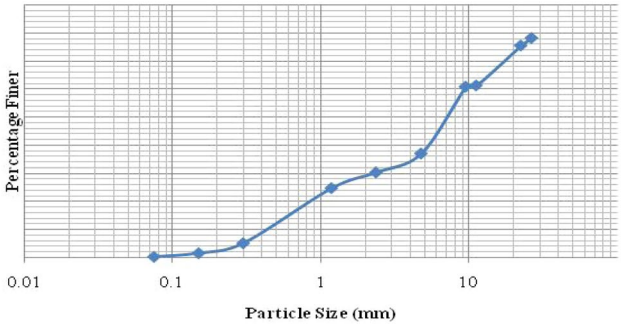


Fig. 2 Particles size distribution for the bed material

Table 1 Seasonal variation of effluent characteristics for phosphate removal by *Phragmites* spp

Season	Parameter	Number of observations (*n*)	Influent		Effluent		Removal efficiency (%)
			Range	Mean ± SD	Range	Mean ± SD	
Autumn	pH	30	7.6–8.0	7.9 ± 0.09	6.7–8.4	7.2 ± 0.45	n/a
	AP (mg/l)	30	1.0–1.3	1.1 ± 0.07	0.2–0.9	0.4 ± 0.28	63.7
	TP (mg/l)	30	2.4–3.3	3.2 ± 0.23	0.8–2.4	1.4 ± 0.51	56.2
Winter	pH	28	6.7–7.6	7.1 ± 0.21	6.6–7.0	6.8 ± 0.10	n/a
	AP (mg/l)	28	1.2–2.7	1.7 ± 0.31	0.1–0.5	0.2 ± 0.11	85.8
	TP (mg/l)	28	2.4–4.7	3.6 ± 0.71	0.3–1.9	0.7 ± 0.36	80.6
Spring	pH	27	7.0–8.3	7.9 ± 0.36	6.8–7.6	7.2 ± 0.32	n/a
	AP (mg/l)	27	1.9–5.6	3.7 ± 1.01	0.1–2.2	0.3 ± 0.40	91.2
	TP (mg/l)	27	6.8–12.5	9.7 ± 1.55	0.1–3.3	0.9 ± 0.88	90.5
	TKN (mg/l)	27	33.6–51.9	46.5 ± 4.29	1.0–18.4	10.6 ± 2.29	76.9
	Fe^{3+}/Fe^{2+}	7	n/a	n/a	0.1–0.2	0.08 ± 0.05	n/a
Summer	pH	16	7.8–8.2	7.9 ± 0.11	7.2–7.3	7.2 ± 0.04	n/a
	AP (mg/l)	16	5.7–17.8	10.9 ± 3.98	0.05–0.4	0.2 ± 0.11	98.5
	TP (mg/l)	16	18.2–23.5	21.3 ± 1.71	0.1–2.6	0.9 ± 0.67	95.7
	TKN (mg/l)	16	98.0–99.4	99.2 ± 0.48	16.8–21.0	15.1 ± 3.20	83.8
	Fe^{3+}/Fe^{2+}	16	n/a	n/a	0.03–0.4	0.09 ± 0.11	n/a

AP available phosphate, *TP* total phosphate, *TKN* total Kjeldhal's nitrogen, *SD* standard deviation

evapotranspiration, favouring significantly high rate of water uptake by the plants resulting in better efficiency of the system during summer season. The removal efficiency was lowest in autumn as the system was fed with only phosphate, as nutrient, and deficiency of nitrates or ammonium ions may result in slow metabolism in plants (Kadlec and Knight 1996). The removal efficiency of TP also followed the same trend as AP. The removal efficiency for TP was maximum (95.7%) in summer season, and minimum (56.2%) during autumn season (Table 1). Total phosphates were present in the system in bound form as well, therefore, the removal of TP was lower than AP. The study concludes that removal of TP was also dependent on the meteorological conditions (ambient temperature and sunshine hours) as shown in Fig. 3. Further, it was observed that an increase in phosphate loading rate increases the removal rate of *Phragmites*-based CW cell, thus highlighting the efficacy of such system at higher influent phosphate concentration (Fig. 4). The CW cell can be particularly useful for decentralised treatment of agricultural runoff which has the possibility of joining a lake or some other water body, thereby causing eutrophication.

3.3 Removal of TKN

The ammonium ions were added to the system externally during spring and summer at the rate of 50 mg/l and 100 mg/l, respectively, since the wetland vegetation was stressed and exhibited partial senescence at the onset of winter. The ammonium ions fed into the system are converted to nitrites (NO_2^-) by microbes, followed by its oxidation to nitrates (NO_3^-) which are taken up by the plants (Ray et al. 2019). There was significant removal of TKN (Table 1) in

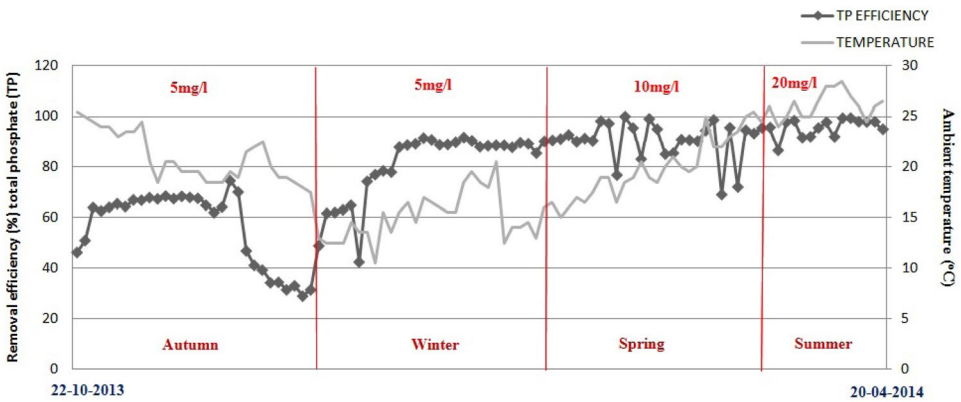


Fig. 3 Seasonal variability of total removal efficiency of phosphate in Phragmites-based constructed wetland

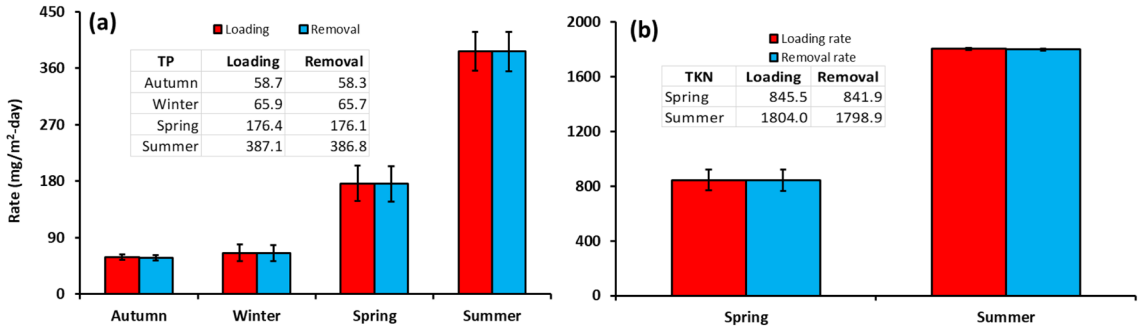


Fig. 4 Variation of mean loading and removal rate (mg/m^2-day) of **a** total phosphate (TP), and **b** TKN for various seasons

the CW system (55.6–90.1%) indicating its assimilation by plants. The average removal efficiency in spring and summer was 76.86% and 83.75%, respectively, indicating that that summer season has the highest removal efficiency which is followed by spring season during the study. The high removal rates of organic nitrogen, ammonia and ammonium from the wastewater show the existence of nitrifying conditions within the CW cell.

3.4 Ferric to ferrous ratio

Ferric to ferrous ion (Fe^{3+}/Fe^{2+}) ratio was analysed to assess the redox conditions, which in turn regulate the availability and removal rate of phosphate. The ratio was observed to be less than unity during the study (Table 1). It is evident that ferrous (Fe^{2+}) ion is present in dominance, thus, representing the dominance of reducing condition in the CW cell. The ferrous ion concentration varied from 0.15 mg/l to 4.1 mg/l, having an average value of 1.34 mg/l, while the ferric ion concentration varied from 0.01 mg/l to 0.14 mg/l with a mean of 0.06 mg/l. This shows that reduction state predominates over the oxidation state in the CW cell. The iron in reduced form associated with substrate media of the wetland cell does not bind with the phosphate which implies that the phosphates are primarily removed by the plants.

4 Statistical analysis

4.1 Mann–Whitney *U* test

Mann–Whitney *U* test was run during the study to ascertain if seasonal variations affect the nutrient removal efficiency or not. Moreover, it was carried out to predict whether the removal is significant or not. For the intended analysis, total phosphate (TP) and total Kjeldahl's nitrogen (TKN) levels during different seasons were used. A number of observations were inferred from this test to assess removal efficiency of these nutrients over the period. It was found that the removal efficiency associated with the variation in seasons are significant ($p < 0.05$) as shown in Table 2. It implies that seasonal parameters such as ambient temperature and exposure to sun played a significant contribution in the nutrient removal.

4.2 Principal component analysis

Principal component analysis was used as a statistical tool for reduction of observed (correlated) data to lesser set of significant free compositive variables. It was used for establishing the inter-relationship between different variables (Ray et al. 2014) viz. ambient temperature (°C), pH, available phosphate removal efficiency (%), total phosphate removal efficiency (%), total Kjeldahl's nitrogen removal efficiency (%) and ferric to ferrous ion (Fe^{3+}/Fe^{2+}) ratio during the study. The rotated component matrix is shown in Fig. 5 in which Component-1 gives an *eigen* value of 2.28. Variables such as temperature (°C), pH, available phosphate removal efficiency (%), total Kjeldahl's nitrogen removal efficiency (%), and ferric to ferrous ion (Fe^{3+}/Fe^{2+}) ratio are present in component-1. This shows a positive correlation, and hence, the increase of parameter of any one variable would make another variable to the increase. The removal of phosphate would primarily be governed by parameters included in component-1, and therefore, this could be

Table 2 Mann–Whitney U test for seasonal comparison of TP and TKN removal (%) in treated wastewater

Seasons	Variables	*P* value
Autumn–winter	TP	0.001
Autumn–spring	TP	0.001
Autumn-summer	TP	0.001
Winter–spring	TP	0.001
Winter–summer	TP	0.001
Spring–summer	TP	0.006
Spring–summer	TKN	0.001

TP- total phosphates; TKN- Total Kjeldhal's Nitrogen

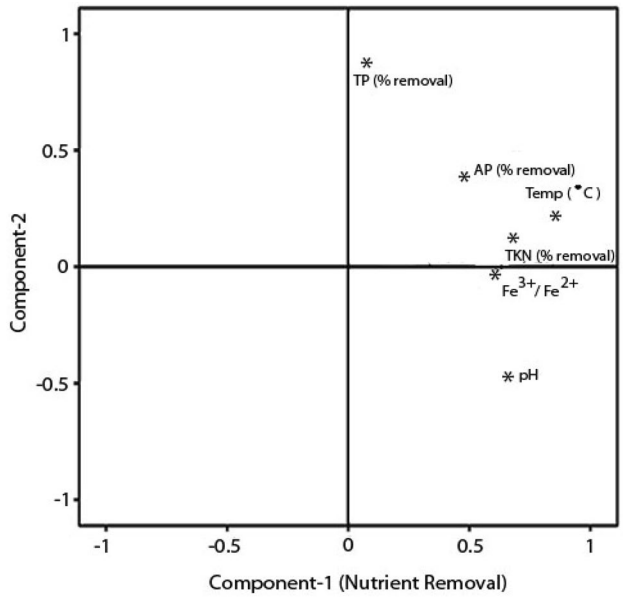


Fig.5 Principal Component Analyses (PCA) plot showing the components in rotated space along with their variables (value > 0.5 is significant)

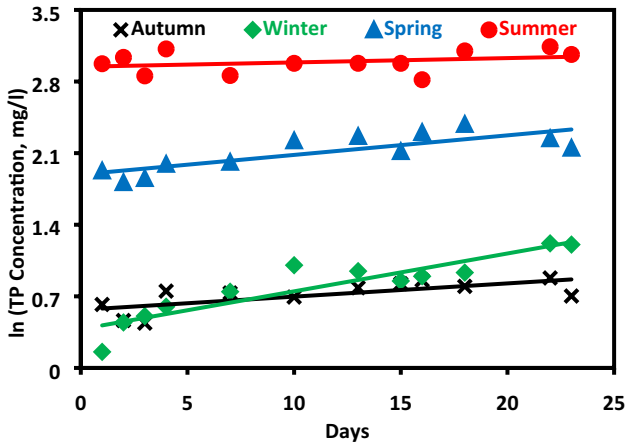


Fig. 6 Plot showing the first order reaction for TP removal for various seasons

called as "nutrient removal". Similarly, component-2 gives an *eigen* value of 1.17 that contains total phosphate removal (%).

4.3 Kinetics of phosphate removal

The order of reaction was determined for the variable concentration of TP, and the first order reaction was considered. For first order reactions, the plot between ln*C* (*C* is concentration) and time gives a straight line and the slope provides the specific reaction rate. The TP removal kinetics was studied for four seasons along with varying initial TP concentration (viz autumn—5 mg/l, winter—5 mg/l, spring—10 mg/l, and summer—20 mg/l). In the present study, the reaction rate followed first order kinetics using the R^2 values as an indication. The corresponding R^2 values for TP removal kinetics obtained were 0.50, 0.83, 0.65, and 0.10 for autumn, winter, spring and summer seasons, respectively (Fig. 6). The R^2 values highlighted that the TP removal kinetics was good and increased for all the three seasons except summer, which indicates that at influent concentration of 20 mg/L of total phosphate, stable removal kinetics was observed for the CW cell used in the present study. The other important factor regulating removal during summer could be temperature and sunshine hours.

5 Conclusion

The study revealed that a *Phragmites*-based constructed wetland can effectively remove phosphate from domestic wastewater, and it can be an effective method for managing eutrophication of water bodies. As *Phragmites* can survive over a wide range of temperature (2–42 °C) without any evidence of stress, it can be a viable resilient option for wastewater treatment at different geographical locations around the globe. Further, it is observed that meteorological conditions play a significant role in nutrient uptake being regulated by rate of evapotranspiration. Higher ambient temperature favours enhanced rate of removal. Apart from it, the plant can treat even high strength wastewater with respect to phosphate (PO_4^{3-} concentration upto 20 mg/l) making it a potential agent to treat even high strength wastewater. As the real wastewater contains nitrogen as a nutrient (NH_4^+ or NO_3^-), it favours growth of *Phragmites* and removal of nitrogen. Simultaneously, the plant has a thick root network which helps maintaining oxidizing conditions in the rhizosphere, although overall environment in engineered wetland may represent clinograde oxygen profile. But, the regions of CW where root penetration is limited or absent may form pockets having reducing conditions. Oxygen to such regions is delivered from rhizosphere by diffusion through the water column, but may not be sufficient to represent high values of ORP, thus resulting in non-binding of Phosphate with Fe(III) of Al(III) ions. Therefore, it is important to maintain slightly reducing and near neutral conditions in rhizosphere to avoid binding of phosphate with the iron and aluminum by maintaining proper depth of substrate in a CW cell. Regular harvesting of above ground tissue is advised for nutrient export, and further use as fodder or raw material for composting, thatching or handicraft, particularly in remote areas.

Acknowledgements The authors acknowledge the help of Mr. S. Nandakumar, M.Tech. Scholar of the Department, throughout this project.

Author contributions HP: conceptualization, execution, data compilation and draft writing; AKH: conceptualization, editing, supervision,

reviewing and editing; KRR: conceptualization, editing, supervision, reviewing and editing.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

Human and animal rights: Authors declare that no involvement of Human and Animals in the research work.

References

Al-Isawi R, Ray S, Scholz M (2017) Comparative study of domestic wastewater treatment by mature vertical flow constructed wetlands and artificial ponds. J Ecol Eng 100:8–18

Binder W, Gottle A, Shuhai D (2017) Ecological restoration of small water courses, experiences from Germany and from projects in Beijing. J Int Soil Water Conserv Res 3(2):141–153

Cameselle C, Gouveia S (2019) Phytoremediation of mixed contaminated soil enhanced with electric current. J Hazard Mater 361:95–102

Haritash AK, Sharma A, Bahel K (2015) The potential of Canna lily for wastewater treatment under Indian conditions. Int J Phytoremed 17(10):999–1004

Haritash AK, Dutta S, Sharma A (2017) Phosphate uptake and translocation in a tropical Canna-based constructed wetland. J Ecol Process 12(6):12

Kadlec RH, Knight RL (1996) Treatment wetlands, 1st edn. CRC Press, Boca Raton

Kajla S, Nagi GK, Kumari R (2021) Microorganisms employed in the removal of contaminants from wastewater of iron and steel industries. Rend Fis Acc Lincei 32:257–272

Metcalf and Eddy (2003) Wastewater engineering treatment and reuses. McGraw Hill, New York

Nandakumar S, Pipil H, Ray S, Haritash AK (2019) Removal of phosphorus and nitrogen from wastewater in *Brachiaria*-based constructed wetland. Chemosphere 233:216–222

Pant HK, Reddy KR, Lemon E (2001) Phosphorus retention capacity of root bed media of sub-surface flow constructed wetlands. J Ecol Eng 17(4):345–355

Ray S, Mohanty A, Mohanty SS, Mishra S, Chaudhary GR (2014) Removal of nitrate and COD from wastewater using denitrification process: kinetics, optimization, and statistical studies. Clean Technol and Environ Policy 16(2):291–301

Ray S, Scholz M, Haritash AK (2019) Kinetics of carbon and nitrogen assimilation by heterotrophic microorganisms during wastewater treatment. Environ Monit Assess 191:451

Schmitt N, Wanko A, Laurent J, Bois P, Molle P, Mose R (2015a) Constructed wetlands treating stormwater from separate sewer networks in a residential Strasbourg urban catchment area: Micropollutant removal and fate. J Environ Chem Eng 3(4 Part A):2816–2824

Schmitt N, Wanko A, Laurent J, Bois P, Molle P, Mose R (2015b) Constructed wetlands treating stormwater from separate sewer networks in a residential Strasbourg urban catchment area: micropollutant removal and fate. J Environ Chem Eng Part A 3(4):2816–2824

Shaltout KH, Galal TM, El-Komi TM (2014) Biomass, nutrients and nutritive value of *Persicaria salicifolia* Willd. in the water courses of Nile Delta. Egypt Rend Fis Acc Lincei 25:167–179

Shan V, Singh SK, Haritash AK (2020) Water crisis in Asian countries: status and future trends. In: Kumar M, Munoz-Arriola F, Furumai H, Chaminda T (eds) Resilience, Response, and risk in water systems-shifting management and natural forcing paradigms. Springer transactions in civil and environmental engineering. Springer, Singapore

Shan V, Singh SK, Haritash AK (2021) Evaluation of water quality and potential metal contamination in ecologically important Bhindawas bird sanctuary. India J Appl Water Sci 11:8

Singh S, Kaushik A, Kaushik CP (2015) Removal of nutrients and organic load from simulated leachate using lab-scale constructed wetland reactors planted with *Canna*. Int J Adv Sci Res 5(1):467–476

Singh P, Bagrania J, Haritash AK (2019) Seasonal behaviour of thermal stratification and trophic status in a sub-tropical palustrine water body. J Appl Water Sci 9:139

Talukdar S, Pal S (2017) Impact of dam on inundation regime of flood plain wetland of punarbhaba river basin of barind tract of Indo-Bangladesh. J Int Soil Water Conserv Res 5(2):109–121

Vymazal J (2005) Horizontal sub-surface flow and hybrid constructed wetlands systems for wastewater treatment. J Ecol Eng 25(5):478–490

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

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

Severity Classification of Diabetic Retinopathy using ShuffleNet

Rohit Gambhir
Dept. of Information Technology
Delhi Technological University
New Delhi, India
rohitgambhir96@gmail.com

Shashank Bhardwaj
Dept. of Information Technology
Delhi Technological University
New Delhi, India
shashankbhardwaj.420@gmail.com

Adwait Kumar
Dept. of Information Technology
Delhi Technological University
New Delhi, India
adwaitkumar555@gmail.com

Ritu Agarwal
Dept. of Information Technology
Delhi Technological University
New Delhi, India
ritu.jeea@gmail.com

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498569

***Abstract*— In India, the number of people living with diabetes is about 72.96 million and large proportion of people suffer from Diabetic Retinopathy which is caused due to elevated level of blood sugar in a diabetic patient that can harm the retina of person and may lead to permanent blindness. The detection of Diabetic Retinopathy is being practiced in such a way that it still needs complex human efforts. The need of the hour is to discover new and better methods that can help further improve identification and treatment of DR (Diabetic Retinopathy). This paper proposes a different method and model to identify and distinguish DR into various severity levels. A novel approach is introduced to detect Diabetic Retinopathy using Shufflenetv2 which is Convolutional neural network. The results are better when compared to other models in literature. Smooth L2 loss function has been used to depict the errors and analyze the results.**

Keywords—Machine Learning, ShuffleNetv2, Resnet50, Diabetic Retinopathy, Loss Function.

I. INTRODUCTION

Diabetes is a chronic disease which harms the blood vessels throughout the body. The eyes begin to get damage when the blood vessels present in retina get blocked due to high blood sugar level and eventually forcing them to leak fluid or bleed. To recover those damaged blood vessels eyes then grow fresh blood vessels which do not function well and may leak blood easily.

Diabetic Retinopathy is a consequence of a bunch of conditions in the eye of a diabetic patient. One of which is identified with mild vascular pain in the retina which ultimately cause the veins and vessels to swell, which further will prompt losing of the human vision.

Eye is one of the vital organs of the human body. Diabetic retinopathy is a concern which directly has a significant impact on the eyes, which could be a consequence of the damage caused to the blood veins, vessels as they are extremely sensitive to the light since retina is at the rear of the eyes.

Despite years of medical and laboratory research, diabetic retinopathy stays the leading reason of vision impairment and blindness amongst operating-age adults, yet the fundamental purpose(s) remains uncertain. The initial phases of diabetic retinopathy generally don't have any indications of the disease. A few groups of patients notice change in their vision, similar to inconvenience in perusing or seeing distant items. In later phases of the illness, veins in the retina begin to bleed into the vitreous humor (gel-like liquid in the eye). On the off chance that this occurs, you may see dim, coasting spots or streaks that resemble spider webs. Now and then, the spots clear up all alone — yet it's imperative to look for treatment immediately. Without treatment, the draining could recur, deteriorate, or cause scarring.

Diabetic Retinopathy may lead to more severe conditions in the patient's eye:

i. Diabetic macular edema (DME): Over a period of time, 50% of the population suffering from DR generally develops DME. DME develops as a result of blood vessels leaking a fluid in the retina which in turn leads to swelling in the macula (part of the retina). Due to the extra fluid present in macula a person suffering from DME may develop blurry vision.

ii. Neovascular glaucoma: DR may lead to the growth of unwanted blood vessels out of the retina which eventually block the fluid flowing out of retina thereby causing a glaucoma.

There are many ongoing projects and research work being done in medical field using the concepts and algorithms of machine learning. One such study is on Covid-19 [1] one of the most widespread disease of 21st century how it affected education system, social life, physical as well as mental health of the people. Different types of deep architecture classifiers have also been used in medical field to contribute in curing mental illness like Alzheimer and suicidal tendencies [2].

The datasets for DR contain the fundus pictures of retina. For research, more than 80 percent of the images present in the dataset were utilized for the training purpose and the remaining were used for testing purpose. We have used various feature extraction methods, some of the features are blood vessels, micro aneurysms, hard exudates. We have used ShuffleNet and have visualized the results in terms of training loss, validation loss and kappa score which are a measure of effectiveness and efficiency required in these types of results in medical field.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:57:29 UTC from IEEE Xplore. Restrictions apply.

The different stages of DR are shown in table I and the results are classified in the similar terms.

TABLE I. TYPES OF DIABETIC RETINOPATHY

Severity level of disease	Observations
No signs of DR	No visual difficulties
Mild non- proliferative DR	Only Micro-aneurysms
Moderate non-proliferative DR	Intermediate state between just micro-aneurysms and severe NPDR
Severe non-proliferative DR	Any one of these: 1. In all 4 quadrants haemorrhage inside retina is greater than 20 2. Bleeding in veins in more than 2 quadrants 3. In more than one quadrant micro vascular abnormalities inside retina 4. No symptoms of proliferative DR
Proliferative DR	One or more of these 1. Neovascularization 2. Vitreous/Pre-retinal haemorrhage

In the first stage (gentle instance of DR), generally some swelling is observed at portions of the blood vessels or the veins that are situated near the retina. The 2nd stage, called as moderate/medium non proliferative retinopathy, chances are that none of the retinal veins will be effected due to the spike in sugar levels of the patient consequently scaling the blockage. In the third stage, genuine or non-proliferative DR conveys with it more impeded veins/vessels, which results in many zones of the retina not being able to receive adequate blood stream again. Subsequently, without suitable circulatory system stream, the retina can't grow more of fresh veins which are required for replacing the damaged veins. In the fourth and the last stage which is known as proliferative retinopathy, is the most severe phase of DR sickness. In this stage of the disease, extra new veins will begin to grow inside the retina, yet they will be exceptionally sensitive and small in structure. Thus, they can spill blood (overflow the blood) which will further incite the vision loss & conceivably bring about the visual impairment.

II. LITERATURE REVIEW

The ever-increasing number of diabetic retinopathy cases generally require uplifting attempts in making advancement to help the investigation of DR. Programmed identification of DR will prevent wasting a lot of investment funds and time and exertion. Wu *et al.* [3] gave a method for programmed identification of microaneurysms in fundus images of retina. They evaluated a network for automated screening of non-proliferative DR and utilized vector machines in the automated determination of non-proliferative DR. A few image pre-processing procedures have been proposed to distinguish DR. Nonetheless, despite of all these previous works, there is still huge scope of improvement in computerized discovery of diabetic retinopathy. In 2015 Jose (4) introduced a web-based system for computerized DR detection. He designed a software program that is completely impartial of user's laptop OS referred to as "computerized DR screening". In this method, sores and red injuries were the two distinct variations from the norm which were observed in results. The clarifications of steps engaged with this work are:

- The starter step in mechanized retinal pathology analysis is image pre-processing. This incorporates different strategies, for example, contrast improvement, closer view/foundation separation, picture de-noising, and then a procedure which is accessible in MatLab Toolbox is utilized in this work.
- The subsequent advance is the element extraction strategy in which appropriate list of capabilities is extricated from the upgraded retinal pictures and from the distinguished anatomical structures[4].

The goal of the feature extraction is two folded:

(a) Generating a list of capabilities which expands the similar classes and limits the class comparability measures.

(b) Aid in dimensionality decrease which at last limits the total time-frame of the classifiers.

Diabetic retinopathy has four phases, in particular, gentle, moderate, serious and proliferative ultimately prompting visual deficiency. The stages are irreversible and the condition can further deteriorate[5]. Consequently, it is significant for early identification to prevent the infection from spreading.

III. FEATURE EXTRACTION METHODS AND ANALYSIS

Image processing can diminish the outstanding burden of screeners as well as generates a focal function for quality confirmation assignments. Along these lines, there has been a development in the utilization of automated picture processing procedures for computerized detection of DR. The accompanying sections show veins, radiates, hemorrhages, microaneurysms and maculopathy identification procedures. These strategies provide the vast majority of the features which are used in computerized DR recognition frameworks.

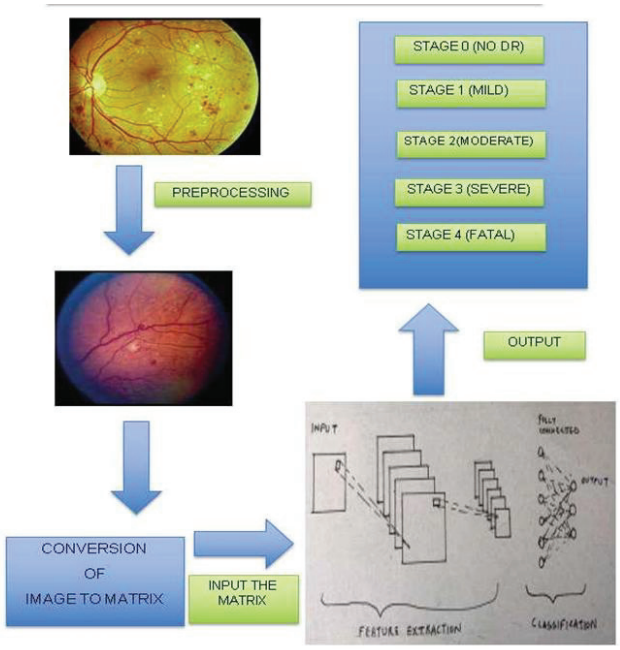


Fig. 1. Proposed System Overview

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:57:29 UTC from IEEE Xplore. Restrictions apply.

The clinical highlights of DR are Micro aneurysms, hemorrhages, exudates and new vessels.

A. Blood Vessel

Computerized fundus photographs of the natural eye give exact form of the veins in the retina. This approach is a brilliant method to observe the eye of the person suffering from DR. Few algorithms were used to reduce the green segment in the RGB fundus images thereby obtaining the exact structure of blood vessels. A two-dimensional coordinated channel was used to identify blood vessels. The idea of coordinated channel was utilized to identify fine details of every segment of blood vessels after their approximation. A novel technique is used to find veins in pictures that divides and enhances nearby vessel due to area-based properties of the natural structure of eye. Scientist have designed a computer aided analysis framework to support doctors in recognizing images of the retina. The irregularities in widths of blood vessels and convergence of veins can be recognized using this framework.

Computerized structure for both extraction and statistical display of the noticeable vascular symptoms from fundus pictures in hypertensive retinopathy was presented. C-means classifier was utilized to find blood vessel points in a cross section. Subsequent to finding the vessels, distinguished portions were associated using previous calculation. Finally, bifurcations and crossing points were perceived, separating vessel end in regards to the vessel structure. Calculations were made using the veins tracker to know the retinal vascular organization caught using a computerized camera.

B. Microaneurysms

In fundus pictures, the micro aneurysms are essential and mini-indications of diabetic retinopathy. Microaneurysms identification is significant, on the grounds that these structures comprise the most conspicuous symptom of DR. The spillage of retinal veins is the primary driver of micro aneurysms, that becomes adjusted red spots on the retina. The primary reports which interfaced these features to diabetic retinopathy were proposed back in 1879. The appearance and vanishing of micro aneurysms are noticed in different periods of fluorescein angiography. In a comparable report both rate of appearing and vanishing of micro aneurysms in early DR were investigated. The micro aneurysms turnover was registered precisely using techniques like the exudate detection algorithm. First of all, the structures in the retina pictures, for example, optic disc and blood vessel tree are to be eliminated. After that a complex arrangement of image processing algorithms were utilized to decide the territories inside the fundus pictures to get micro aneurysms.

C. Hard Exudates

One of the early indications of DR is the recognition of exudates in the retina. In general, there are various kinds of exudates, in particular hard exudates and soft exudates. Soft exudates (cotton fleece spots) are similar to gentle yellow or white areas with distracted edges whereas hard exudates are exemplified as yellow waxy patches inside the retina.

Identification of exudates is a significant trademark for analysis and the clear form of them causes us to remember them. To identify hard exudates, the red part present in CMY picture is extricated and a bounded binarization determined with respect to the standard deviation of the red segment is implemented. For further improving the binarized picture the retinal structure is changed to white tone and optical circle veil is added. At that point, a widening activity is applied to the subsequent picture and enormous ceaseless areas are eliminated. At last, a disintegration activity is executed prior to processing the thickness of hard exudates.

IV. DETECTION PROCESS OF DR

Blindness detection framework is innovation that can be executed in existing lab and get the outcome without the assistance of profoundly qualified specialists. The fundamental advantage of the framework is that we might detect the beginning phase of diabetic retinopathy.

A. Pre-processing

Computerized severity detection of DR (SDDR) is introduced within the current studies via bag of functions (BoF) approach. It is an adaptive technique to represent the photograph in a sturdy manner and categorize them. Adaptiveness of BoF(bag of functions) is one of the advantages because it allows picture collection to be processed and also identifies visible styles of the whole image collection. The approach targets five severity tiers of DR:

0 - No DR: Normal eye without any indications of illness

1 – Mild: Microaneurysms (growing of veins), show up at this stage.

2 – Moderate: Blood vessels in retina swell up

3 – Severe: Blood vessels are hindered, impeding the blood flow to the retina.

4 - Proliferative DR: Many veins in the retina stop, don't give adequate blood stream because of which it develops new strange veins.

As training a model with large size data will require a lot of resources and time, so to maintain the size of data we need to reduce the size of images to 256*256.Along with reducing the size of images, further pre-processing was required in order to make the images suitable for training model. Gaussian Blur Subtraction: it is the outcome of blurring an image upon using a function known as gaussian function. It is a mainly brought in use to decrease image clamoring and prevent the details from diminishing. The outcome of Gaussian blur is as if the image is seen via a translucent glass. Gaussian blur subtraction is required in order to enhance the image quality so that a higher number of features could be extracted from each image.

B. Training and Testing

Training is the significant period of planning a deep learning framework where we arrange the model using researched dataset. The structures we utilized for preparing the model are ShuffleNet V2, Resnet50, DenseNet121, DenseNet169, Densenet256. Resnet50 architecture performs training using 48 convolutional layer and also incorporates one average pool layer and one Max pool layer 1 whereas ShuffleNet V2 incorporates 162 convolutional layers for training purposes thus it is considered superior in terms of results over other previously used models such as SENet and ResNet.

At first the loads were used without utilizing top layer and planned last layer. Cluster size of 32 is used to prepare

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:57:29 UTC from IEEE Xplore. Restrictions apply.

the model and image of size 256*256 is used as input. We utilized optimization technique that have less training cost contrasted with proceedings of the Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020). Thus, in order to improve the efficiency of training the model Adam optimizer algorithm was used. Also, prepared the model with 7 epochs. Customary feedforward network that utilizes nth layer yield to $(n+1)^{th}$ layer's input was used where ShuffleNet add a leap association that sidestep the non-linear transform.

V. RESULTS AND EXPERIMENTS

A train dataset is created by using Create Dataset function and by applying transforms to the dataset. In this we split the training and test set in 85:15 ratio. The approach of splitting the dataset in the 90:10 ratio and 80:20 ratio for further checking and understanding the working of the dataset was also used. Here, pre-trained models were used.: Resnet50 and ShuffleNet V2 after some study and observing that they work well on the dataset that is available to us in terms of image quality and level of images that are available from the processing and feature extraction of the fundus images of the eyes. Here, multiple layers in the model were used to improve the val_kappa score and to reduce the val_loss. Epoch refers to number of iterations of training dataset completed by the architectural model. Out of 7 epochs that were used to prepare model best results were obtained for 4th epoch (epoch number 3 as indexed from 0).

Two loss functions were also used in order to minimize the error: nn.SmoothL1Loss() also known as LAD which stands for Least absolute Deviations and nn.SmoothL2Loss() also known as LS which stands for Least Square errors.

$$L1 = \sum_{i=1}^{n} |y_{true} - y_{predicted}|$$

$$L2 = \sum_{i=1}^{n} (y_{true} - y_{predicted})^2$$

where y_{true}: actual value

$y_{predicted}$: output value of model

The results in case of the latter were better than the other loss function and thus this loss function was used for further training. Different learning rates and number of epochs were used to maximize the kappa score and to reduce the validation loss. We have taken kappa score as measure for comparing the results as it is a norm that kappa score is a better measure as compared to normal percentage settlement calculation, as it considers the probability of agreement happening accidently. For better results the model with higher kappa score is better which is showcased by ShuffleNet V2 with higher score in comparison to other models in this case and in terms of loss, lower the loss better will be the model for practical usage.

A. Visualization of Training and Validation Loss for different epochs

Training loss refers to error obtained on training data on comparison of actual/true value with the value predicted by the model used for training. Validation loss refers to error obtained on the validation dataset when the trained model is tested over validation data. Both training and validation loss showed similar trend as both decreases with increase in number of epochs to a certain point.

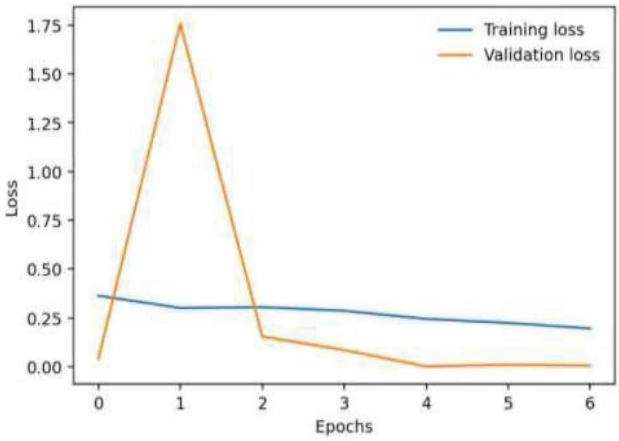


Fig. 2. (ShuffleNet V2)

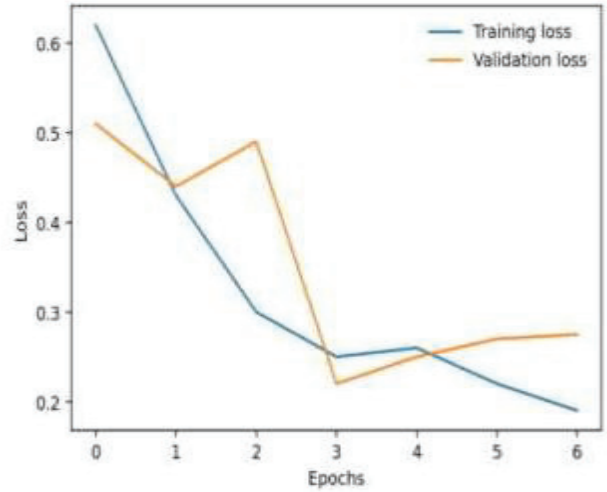


Fig. 3. RESNET 50)

B. Visualization of Kappa Score for different epochs:

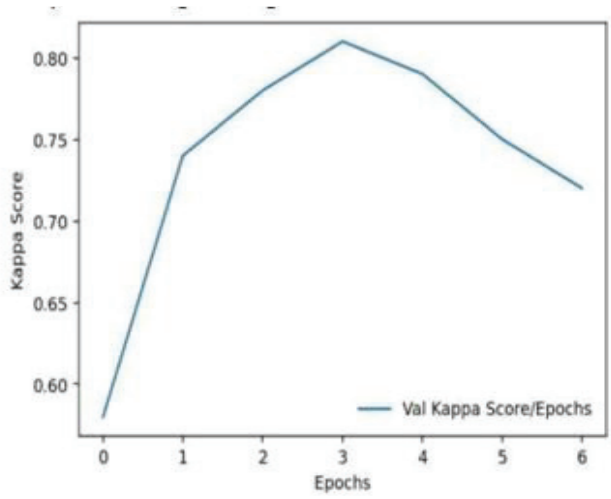


Fig. 4. ShuffleNet V2

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:57:29 UTC from IEEE Xplore. Restrictions apply.

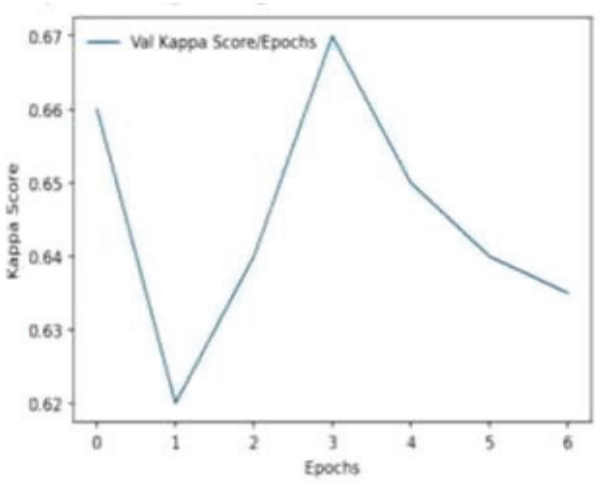


Fig. 5. RESNET 50)

VI. CONCLUSION

In the presented research, a framework is derived for distinguishing the severity levels of DR into 5 distinct classes using visual features and a model architecture ShuffleNet V2, to get an effective system. The main emphasis is on identifying the severity class to which the image belongs, upon applying preprocessing on the input images. It is observed from the literature that there have been many research work and studies but analyzed severity levels of DR are not that extensive. More emphasis was given to the identification of the effects of DR. This type of automated system should be used with high accuracy, high sensitivity and good percent of specificity. The performance in these types of experiments can be further improved with increase in diversity of data and improved features.

REFERENCES

[1] COVID-19 and its impact on education, social life and mental health of students: A survey Chaturvedi K, Vishwakarma D, Singh N Children and Youth Services Review (2021)

[2] V. Singh and K. Sharma, "Empirical Analysis of Shallow and Deep Architecture Classifiers on Emotion Recognition from Speech," 2019 6th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/ 2019 5th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom), Paris, France, 2019, pp. 69-73, doi: 10.1109/CSCloud/EdgeCom.2019.00-16

[3] Dai B, Wu X, Bu W (2016) Retinal Microaneurysms Detection Using Gradient Vector Analysis and Class Imbalance Classification. PLOS ONE 11(8): e0161556.

[4] Jose Tomas Arenas Cavalli, Sebastian A.Rios, Mariano pola and Rodrigo Donoso "A web-based platform for Automated DR Screening," International 37 conference on intelligent information and knowledge based and engineering systems, pp. 557.–.563, 2015.

[5] Sharath kumar. P. N, Deepak. R. U, Anuja sathar, Sahasranamam. V, Rajesh kumar R. "Automated detection system for DR using two-field fundus photography,".international conference on advances in computing Communications(ICACC), pp. 486–494, 2016

[6] U R Acharya, C. M. Lim, E.Y.K. Ng, C Chee, T. Tamura "Computer based detection of D.R stages using digital fundus images.," Journal of engineering in medicine, pp. 545–553,2009.

[7] Kayte, Sangramsing. (2015). Review of Automated Detection for Diabetes Retinopathy Using Fundus Images. International Journal of Advanced Research in Computer Science and Software Engineering. 2.05. 8.

[8] Automatic Detection of Diabetic Retinopathy: A Review on Datasets, Methods and Evaluation Metrics

[9] Detection of severity level of diabetic retinopathy using Bag of features model - IET Journals & Magazine

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:57:29 UTC from IEEE Xplore. Restrictions apply.

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 08 | Aug 2021 www.irjet.net p-ISSN: 2395-0072

Solar Air Process Heating Systems for Drying of Agricultural Products

Rohit Kumar[1], Suraj Kumar[1], Sanchit Sharma[1], Dr. J.P. Kesari[2]

[1]Student of B.Tech in Mechanical Engineering, Delhi Technological University, Delhi, India
[2]Professor of Mechanical Engineering, Delhi Technological University, Delhi, India

---

Abstract – *Solar drying is one of the oldest and easiest ways of drying known to us. The purpose of this work is to design, develop and evaluate the indirect performance solar dryer prototype in a passive mode for drying of agriculture products like fruit & vegetables. Drying is one of the oldest ways to use solar energy where products such as vegetables, fruits, fish and meat to come in direct contact with the sun. This method has many disadvantages such as spoilage products caused by rain, wind, dust, insect infections, animal attacks and fungi. Foods must dry fast, but will harden out due to drying speed before the moisture has a chance to evaporate inside and it will affect the quality of dry products due to over-drying. Design of an indirect solar dryer system minimize these losses. The whole designing of 3-D indirect solar dryer has been done with the help of Google SketchUp. In this review, we also attached a moisture collector for the smooth extraction of any remnant moisture. This design of indirect mode prototype is employed in a graphical method, after finding out the parameters such as temperature, moisture contents and weight loss. This type of solar dryers has good performance and provide better products.*

Keywords: **– Solar Dryer, Design Concept, Modelling Methodology, Analytic Study, Future Availability, Renewable Energy, Environmental Friendliness.**

1. INTRODUCTION

Many agricultural products have a high moisture content which, due to biochemical reactions, makes them vulnerable to microbial and other spoilage. Therefore, to minimize the moisture content of the products, drying or dehydration activities must be carried out as preventive steps. One of the key problems faced by developing countries is the preservation of agricultural products. Over a long time, the rising nutritional requirements of the ever-expanding populace of these nations will deteriorate these issues.

Major amounts of products are being ruined in many developing nations because of unstable infrastructure &, improper processing facilities. It should be noted that around 70 percent of foods are being spoiled, due to traditional drying methods, especially in rural places. Drying these products will help solve these issues, while also making a major contribution to improving the income and supply situation of the population.

Drying is an important method of food preservation that is often carried out at farm level immediately after harvest or, in particular. The thermal drying of vegetables, fruits and meat allows for longer storage periods and easier transport. The food is placed in boxes with a transparent lid with direct sun dryers. In addition, because of the greenhouse effect, the temperature in the dryer is lifted and the air circulation is regulated by vents. In indirect sun dryers, food is not exposed to direct sunlight as fresh air is heated separately from the food chamber. For drying foods that lose nutritional value when exposed to direct sunlight, this approach is preferable. In contrast to current drying processes, solar drying can only be effective when it shows tangible benefits. Solar dryers avoid degradation of goods by pollen, insects, etc. as opposed to the conventional method of drying outdoors in an open area, thereby guaranteeing consistency. They help small-scale farmers to convert their harvest into products that can be processed and exchanged, which can be sold off-season at higher prices. The constant temperature and ventilation allow the drying process to be stable, resulting in increased product quality and higher prices. Investment costs for solar dryers, however, differ considerably depending on the scale of the solar dryer, the locally available components and the environmental factors, such as slope and side exposure, and the rainy seasons.

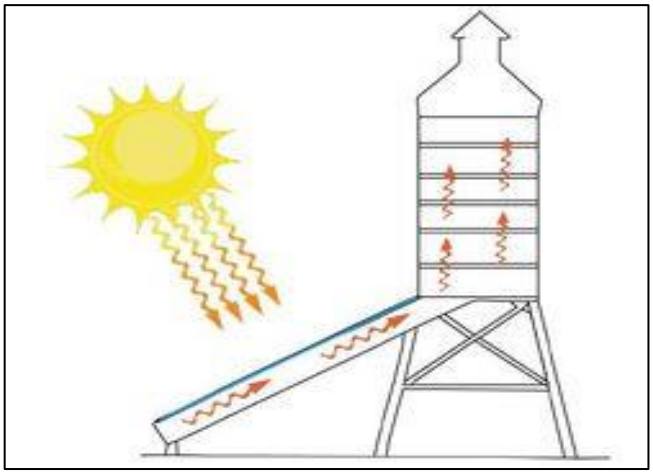


Fig -1: Schematic view of a solar dryer
Source: Energy Pedia

Drying successfully relies on:

- Enough heat, without cooking the food, to draw out moisture;
- Dry air to absorb the moisture released; and
- Sufficient air ventilation to take moisture down.

The drying of agricultural goods for improved preservation, long-term storage and export is a big part of the food processing industry. A big chunk of the drying is still achieved in India by drying the product out in the light. This not only takes a lot of time (usually in days) but is often very unhygienic because of possible environmental damage. It is possible to use a solar dryer for fruits (apples, apricots, grapes, pineapples and bananas, melons, plums, beets, mangoes, dates, figs), vegetables (cabbage, broccoli, peppers, basil, onions, squash, tomatoes, asparagus, celery, potatoes, peas, carrots, cassava, peppers, yams, red cedar, mahogany) corn, maize, rise, cassava, cocoa, fish, meat, mushrooms, spices (dried chili peppers, garlic) tea, coffee, cacao, tobacco, cashew and macadamia, milk, hay, copra (kernel of the coconut), for clothing also like, wool, kindling and as well as for treating timber.

The use of solar dryers is a more productive way; they increase the rate of removal of moisture, thereby accelerating the drying process while reducing the chances of contamination. The use of the industrial solar dryer thus decreases the drying time and helps you to increase the amount of output.

However, it's best to opt for gas fired dryers for high production capacity, as solar dryers need a lot of space. The drying time is significantly reduced by this gas-fired dryer and thus allows high output rates. These are not standard dryers and must be custom designed according to the requirement, which is typically based on the moisture content to be dried in the product, the necessary rate of moisture removal, total output per day, etc. With the preservation of original colour and taste, solar dryer eliminates excess moisture from the product as well as maintains the product clean of any environmental influences, dust, dirt, animal or bird droppings. [1]

1.1. Solar Dryer Can Be of Two Types

1. Direct Gain Dryer
2. Indirect Gain Dryer

In their way of using solar energy, the two are distinct from each other. In direct solar drying systems (fig.2), the products are directly affected by solar radiation. The goods are heated directly by sunlight and moisture is released by the outlet into the air (usually a chimney). Whereas the indirect drying system (fig.3) use solar radiation to heat air inside the heater chamber, which is then moved to the dryer chamber where the hot air passes through the products, removing their moisture content along the way and leaves through the chimney at the top of the drying chamber.

An indirect solar dryer system has two main components, i.e., solar air heater and dryer cabinet. The air heater is a box like a structure with a glass cover at the top to allow solar gain to enter (usually wooden or metal). Inside the heater, an absorber plate is located that absorbs heat from solar energy and transmits it to the surrounding air, heating it. In general, the absorber plate is a black sheet of painted metal that may have fins or may be corrugated to improve the absorber's surface area. The heated air is moved to the dryer cabinet, which is again a cabinet with perforated trays (wire meshes) and a chimney at the top to allow the air to escape to the outside as a framework made of wood or metal. [2]

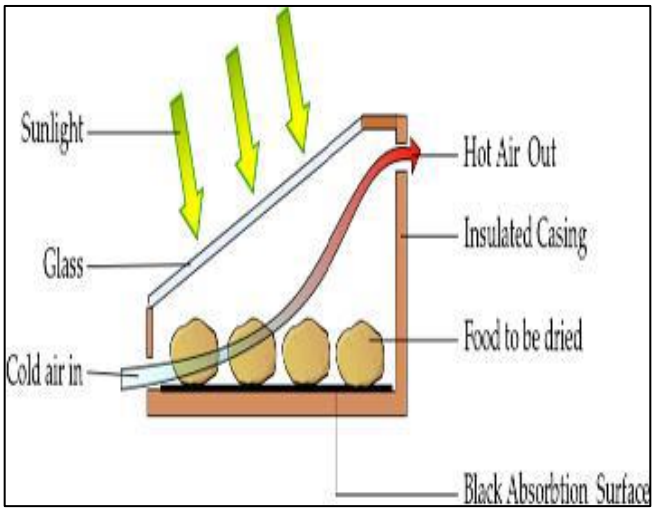


Fig -2: Direct Dryer
Source: Wikipedia

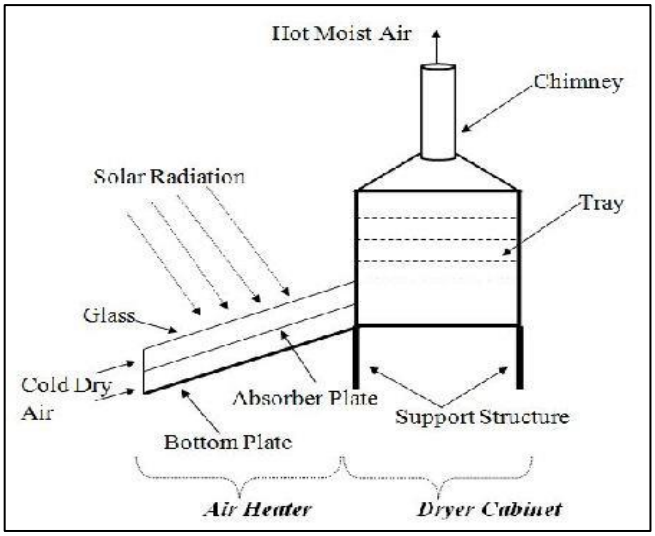


Fig -3: Indirect Dryer
Source: Wikipedia

1.2. Quality Changes During Drying

The act of applying heat to dry material does not simply remove moisture. But dry can affect product quality. These effects differ only from those Commonly encountered events that will be described here.

1. Browning: - It is the dissolution of material during drying, which can either be caused by Physical processes or chemical reactions. It depends on the combustion effect Time and temperature on the moisture of drying material.
2. Case Hardening: - it is characterized by the surface drying of the material and being relatively impermeable. For the further flow of moisture, but with some interior remaining at high humidity material. But with the drying of most vegetables and fruits, the matter does not harden.
3. Rehydration: - It is not that dehydration is a complete reversal of the dehydration process. As in the case of the production of a similarly rehydrated product, rehydration is basically not so important because the dried product is used by adding it to a soup or stove directly near the cook.

1.3. Classification of Industrial Dryers

A wide assortment of dryer plans has advanced throughout the years in terms of working the drying unit. The different method has been suggested as the best way to define them but before seriously considering the methods of classification, it is necessary to define, but Broad principles,

I. How thermal vital prerequisites are provided.
II. How to operate the dryer

In the first place, the heat must be transferred to the wet material for drying operation. heat can be applied in at least one of the accompanying ways.

(a) Convection whereby the heat medium where air or combustion product is in direct contact with wet material.
(b) Conduction, where heat exposure occurs indirectly through contact with the wet material.
(c) Radiation, where heat is transmitted directly and completely to a hot body, a material wet by heat radiation. [3]

1.4. The Following is a Comparison of Each Type Industrial Drying with Illustrations of Typical Application

1) *Conduction VS Convection Dryers*

A distinction can easily be made between conduction and convection drying. Relatively large solid gas separating equipment is often required for convection dryers (e.g., flash and spray dryers). With the increase in inlet temperature of the drying gas, the thermal efficiency of convection dryers also increases. While within conduction dryers, there is no such effect. The heat energy of the drying medium in convection drying is transmitted by convection to wet materials which have direct contact with hot airflow. As a drying medium, the hot air is both a heat carrier and a wet carrier in convection drying. In conduction drying, the heat energy is conveyed in terms of conduction to wet materials. Wet products do not have direct contact with heating media.

2) *Radiation VS Convection Dryers*

In the case of radiation drying, heat energy is produced by the emission of infrared radiation sources in the form of electromagnetic waves projected for water evaporation on the material surface. This process is often called infrared drying, and it is possible to use electricity and coal as the source of infrared radiation. Radiation drying is much quicker and requires less energy than convection drying, since it only heats the surfaces to be dried. However, it is only possible to dry up surfaces exposed to radiation and this restricts the possible form of the object Another downside is that it does not heat various colors, pigments and surface structures to the same degree. Convection drying does not have these limitations on goods, but it is important to heat a large thermal mass. Typically, drying happens in two separate ways. Usually, radiation drying is carried out with fair energy savings and 50 percent-80 percent of the energy is used for evaporation. The energy usage of convection drying is always low and it is pretty common to use just 10%-20% of the energy for evaporation. [3,4]

2. PROBLEM IDENTIFICATION

The solar air dryer is an old concept, but in the modern world, there are many different parameters considered during its construction. Advantages as well as disadvantages such as friction loss, vibration, expansion of acrylic.

Friction losses on glass, cross-section and reducing leakage due to heat and airflow it can't be identified that air in small parts. After considering all the factors, we have taken the decision to make solar dryers for agricultural products, which is less expensive, is more efficient for poor people or farmers.

3. DESIGN AND METHODOLOGY

3.1. Introduction to Google Sketchup Software

SketchUp is a 3D modelling computer program application created by Trimble Inc. A wide scope of drawing applications, for example, engineering, interior design, landscape design, civil and mechanical designing, film and video game design. [5]

How Google SketchUp is different from other 3D designing software: -

- SketchUp is much easier to used and master.
- SketchUp has a good and simple learning curve. you need to only start and create the modelling is a simple object to see.
- The user interface is easier to navigate and it's also a friendly interface.
- SketchUp is a 3-dimensional modelling software where you can draw and edit any 2D and 3D model with tools. push and pull are a tool that is used to convert any flat surface to create into 3D shapes.
- SketchUp has a warehouse option which is used to already created model available for anyone access it.

3.2. Design Considerations & Assumptions

Table -1: Parameters for Solar Dryer

S.No.	Parameters	Description
1.	Location	28°42'41.3"N 77°15'44.8"E
2.	Drying Period	December
3.	Solar Irradiation for December	15.3MJ/m2/hr
4.	Food Product	Agriculture Products (Potato & Carrots)
5.	Mode of Heating	Indirect
6.	No. of Glazing	1
7.	Glazing Material	Glass
8.	Loading Provision	Door at the Sideway of the Cabinet
9.	Air Outlet Provision	Air Conduct Passage through the Upper Most Section of the Back
10.	Number of Trays	4
11.	Air Circulation Mode	Natural Convection
12.	Drying Capacity	10 kg
13.	Thickness of Plantain Fillets	3 mm
14.	Construction Materials	Wood, Glass, Aluminium Sheet
15.	Insulation Used	Glass Wool
16.	Thickness of Aluminium Sheet	3 mm
17.	Thickness of Glass	5 mm

3.3. Design

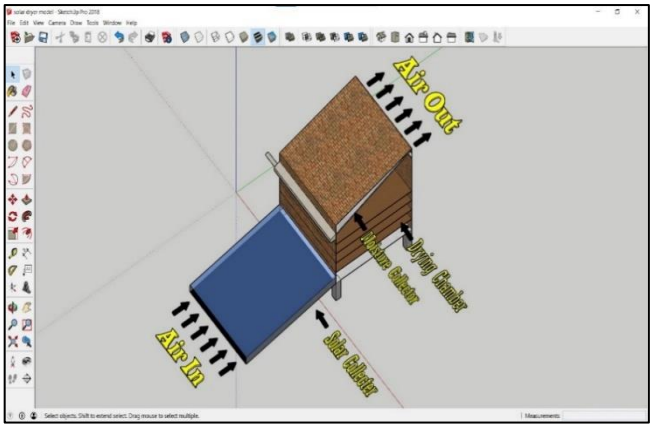


Fig -4: Isometric Model of Solar Dryer

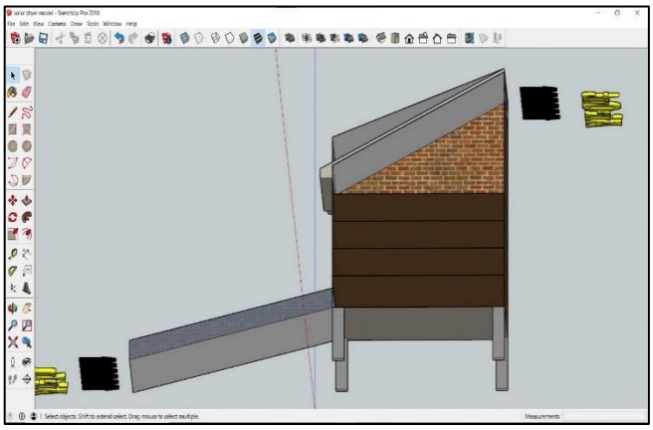


Fig -5: Top View of Model of Solar Dryer

The solar collector of a solar air heater has a box that is made of plywood on the bottom, where it acts as fiber glass on its top and body. Solar dryer is work on simple concept. then basic principles of solar dryer are:

- Converting light into heat energy: The plywood is painted black on one side. This helps in increasing the capability of plywood for converting light into heat.
- Trapping heat: It is of great importance to separate the air inside and outside of the dryer. For this, we used solid fiberglass as a cover. Once the light converts into heat, the glass fiber traps the heat inside in the dryer.
- Interaction of heat and product: There are two methods of the dryer, forced convection dryer and natural convection dryer, both of dryer transfer the heat in food. in this project, natural convection method is using whereas there is a passive type of solar dryer.
- Moisture collector: We added a new concept of removing of evaporated moisture from the drying

chamber with the help of attached incline aluminium sheet. [6]

4. COST ESTIMATION

Table -2: Bill of Materials [7,8,9]

Part Name	Quantity	Unit Cost	Total Cost (Rs.)
Aluminium Sheet	1	1225	1225
Angle Iron	4 (each 45 cm)	3/ cm	540
Wire Gauge	4	120	480
Plywood	135.225 sq. ft.	46.63/ sq. ft.	6305
Blank Paint	1 (500ml)	149	149
Fiber Glass	11.725 sq. ft.	25/ sq. ft.	300
Miscellaneous			500
TOTAL			**9499**

5. EXPERIMENTAL SETUP

5.1. Design Implementation

The ambient temperature of the solar dryer during the direction was determined in writing. After physical experiments with the help of a calibrated thermometer, in which in this real-time project, we offer modern design, compact construction and large display of an indirect solar dryer to conserve useful food Products. Thus, this solar dryer ultimately shows sufficient capacity to ensure rapid clean up to a safe moisture level safely and even more clearly. Improved quality of practical dried product.



Fig -6: Front View of the Indirect Solar Dryer Prototype



Fig -7: Top View of the Indirect Solar Dryer Prototype

5.2. Observations

The individual fact of moisture is wiped out during drying in the month of December 2020, both outside and inside the assembly, as shown below. During the drying period, room temperature is 21OC and compared percentage removal of moisture in preserved food with the help of the solar dryer. the atmospheric air (which is present in the surrounding) following table shows based on observational data. We also took various products for experimental calculation. Average the efficiency of the dryer in a day which was satisfactorily 15% with the rate of removal of moisture of various samples like potato, carrot was found to be 64%, 58% respectively. All

practical offenses were on the underlying support of a day completely sunny. [10]

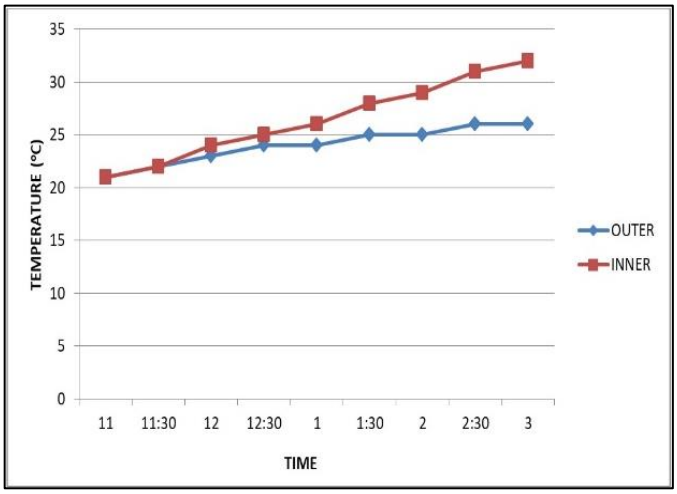
The coming after data is given below:

Table -3: Day-wise Temperature and Weight Measurement

DAY	Date	Time	Outer Temp.	Inner Temp.	Weight at Morning		Weight at Evening	
					Potato	Carrot	Potato	Carrot
1	06-12-2020	11:00 am	21°C	21°C	500 g	500 g		
		11:30 am	22°C	22°C				
		12:00 pm	23°C	24°C				
		12:30 pm	24°C	25°C				
		01:00 pm	24°C	26°C				
		01:30 pm	25°C	28°C				
		02:00 pm	25°C	29°C				
		02:30 pm	26°C	31°C				
		03:00 pm	26°C	32°C			219 g	173 g
2	07-12-2020	11:00 am	21°C	21°C	219 g	173 g		
		11:30 am	21°C	23°C				
		12:00 pm	23°C	25°C				
		12:30 pm	24°C	26°C				
		01:00 pm	24°C	27°C				
		01:30 pm	25°C	28°C				
		02:00 pm	26°C	29°C				
		02:30 pm	26°C	30°C				
		03:00 pm	26°C	30°C			136 g	118 g

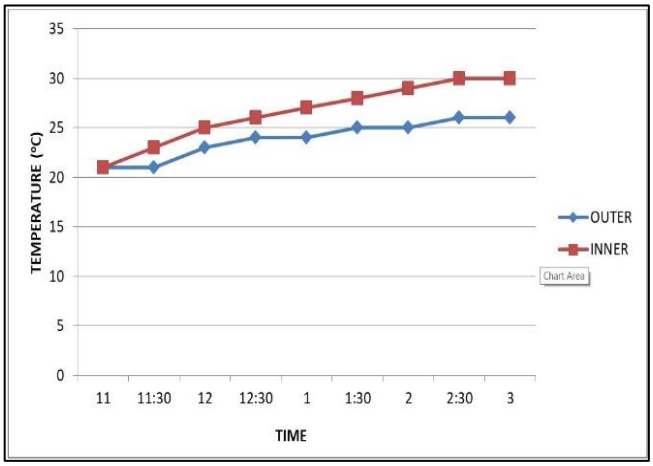
- **Graphical Representation of Temperature Vs. Time of Drying Potatoes and Carrots:**

DAY 1



Graph 1: Day 1, Temp Vs. Time graph of Potatoes and Carrots

DAY 2



Graph 2: Day 2, Temp Vs. Time graph of Potatoes and Carrots

- **Original Appearance of The Potatoes and Carrots Before and After 4 hours of Drying in a Shiny Day:**



Fig -8: Potatoes and Carrots Before Drying



Fig -9: Potatoes and Carrots After Drying

6. CALCULATION

6.1. Potatoes

Moisture content of potato before drying = 80%,
Weight before processing = 500g,
Weight after processing = 136g,
Amount of moisture content = $500 \times \frac{80}{100} = 400g$
Amount of moisture removed = $500 - 136 = 364g$
Moisture remains after drying = $400 - 364 = 36g$
Moisture content after dried = $\frac{36}{136} \times 100 = 26.47\%$

6.2. Carrots

Moisture content of carrot before drying = 88%,
Weight before processing = 500g,
Weight after processing = 118g,
Amount of moisture content = $500 \times \frac{88}{100} = 440g$
Amount of moisture removed = $500 - 118 = 382g$
Moisture remains after drying = $440 - 382 = 58g$
Moisture content after dried = $\frac{58}{118} \times 100 = 49.15\%$

7. RESULTS AND DISCUSSION

The general trend in the temperature profiles obtained indicates a rise from 11:00 to 15:00 and a decrease thereafter (Graph 1 and 2). It was noted that the temperature of the solar collector was always higher than the heating chamber. For the cabinet, maximum temperatures of 32oC and 30oC from day 1 and day 2 respectively were recorded. On day 1, the amount of water content removed from the product was higher than day 2. Also, the amount of moisture content removed within the products was observed to be proportional to the temperature of the heating chamber, hence the maximum removal occurred when the chamber temperature was between the third and fourth hours. The amount of moisture content extracted from potatoes is 281g & 83g for day 1 & day 2. For carrots, on day 1 it is around 327g & on day 2 it is 55g. For the span of 8 hours the percentage of moisture extracted for potatoes is 26.47 percent and 49.15 percent for carrots. [11]

8. CONCLUSION

Solar dryer works operated on solar energy. As in nation Out of 365 days in India, 300 days of sunlight a very valuable thing of solar energy. The government is encouraging toward the use of renewable energy as a solar panel, solar pump, solar heater, solar light, etc. and the government provides subsidies with the condition. Therefore, it is a good chance to take advantage of their Plans. We don't need anything from the cost of solar power, just a solar panel is needed; Sunlight is free and is also a non-ending source of energy.

REFERENCES

[1] Solar Energy. (2019, June 17). Energy Pedia. Retrieved December 29, 2020, from energypedia. info:https://energypedia.info/wiki/Solar_Drying

[2] Sharma, A., & Mundari, N. D. (2017, June). A Review on Solar Air Heating and Drying Techniques. International Journal of Engineering Technology, Management and Applied Sciences, 5(6), 424-430. doi:10.13140/RG.2.2.24453.76005

[3] JM, C. (1979). newengineeringpractice.blogspot, 5. (O. Butterworth, Editor) Retrieved January 5, 2021, from newengineeringpractice.blogspot.com: http:// newengineeringpractice.blogspot.com/2012/01/solar-dryer.html

[4] yhdrying. (2017, June 26). yhdrying. Retrieved January 7, 2021, from www.yhdrying.com: http://www.yhdrying.com/dryer/The-Conduction-Drying-Methods-of-Dryer.html

[5] Wikipedia. (2021, January 2). Wikipedia. (T. Inc., Producer, & SketchUp) Retrieved January 10, 2021, from wikipedia.org: https://en.wikipedia.org/wiki/SketchUp

[6] Network, C. T. (n.d.). Climate Technology Centre & Network. (K. Larsen, Editor, & Denmark) Retrieved January 24, 2021, from www.ctc-n.org: https://www.ctc-n.org/technology-library/renewable- energy/solar-dryer

[7] Amazon (n.d.). ASIAN PAINT OIL PAINT IN BLACK COLOUR SIZE 500 ML

[8] Indiamart. (n.d.). Wire Gauze - Wholesaler & Wholesale Dealers in India

[9] Indiamart. (n.d.). L Shaped Slotted Angle, 2.50,3.00 Mm, 2.5, 80 Mm X 40 Mm

[10] Babu, K. N., Kumar, P. S., & Yamuna, D. (2019, July). A Research on Development of a Fixed Solar Dryer with a Practical Research. International Journal of Innovative Technology and Exploring Engineering, 8(9S2), 820-824. Retrieved December 28, 2020, from https://www.ijitee.org/wp-

content/uploads/papers /v8i9S2/I11680789S219.pdf

[11] Adelaja, A. O., & Babatope, B. I. (2013). Analysis and Testing of a Natural Convection Solar Dryer for the Tropics. Journal of Energy, 2013. doi:https://doi.org/10.1155/2013/479894

BIOGRAPHIES



ROHIT KUMAR, (B.Tech in Mechanical Engineering, Delhi Technological University)



SURAJ KUMAR, (B.Tech in Mechanical Engineering, Delhi Technological University)



SANCHIT SHARMA, (B.Tech in Mechanical Engineering, Delhi Technological University)

Journal of Fluorescence
https://doi.org/10.1007/s10895-021-02788-z

ORIGINAL ARTICLE

Structural, Electronic and NLO Properties of 6-aminoquinoline: A DFT/TD-DFT Study

Nupur Pandey[1] · Mohan Singh Mehata[2] · Sanjay Pant[1] · Neeraj Tewari[3]

Received: 11 May 2021 / Accepted: 13 July 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract
A computational study based on the DFT/TD-DFT approach was performed to explore various properties of 6-aminoquinoline (6AQ). The geometrical parameters, molecular orbitals (MOs), electronic spectra, electrostatic potential, molecular surface, reactivity parameters and thermodynamic properties of 6AQ were explored. The absorption and emission spectra of 6AQ in solvents have been estimated by TD-DFT coupled with the PCM model and correlated with the available experimental results. Depending on the solvents, the computed absorption maxima of 6AQ were noticed between 327 nm – 340 nm and ascribed to $S_0 \rightarrow S_1$ transition. The simulated emission maxima were obtained between 389 to 407 nm and ascribed to $S_1 \rightarrow S_0$ transition. On increasing the solvent polarity, both the emission and absorption maxima showed a bathochromic shift. The LUMO and HOMO were localized on the entire molecule. It was observed that the lowest excited state is possibly the $\pi \rightarrow \pi^*$ charge-transfer (CT) state. The natural bonding orbital (NBO) study points out that ICT plays a significant role in stabilizing the molecular system. Moreover, the NLO (nonlinear optical) properties (polarizability, first-order hyperpolarizability and dipole moment) were computed using different hybrid functionals. The estimated values indicate that 6AQ can be considered a desirable molecule for further studies of the NLO applications.

Keywords 6AQ · DFT/TD-DFT · ICT · Solvent polarity · Electronic spectra · NLO properties

Introduction

Quinoline derivatives have been of particular concern in medicinal and organic chemistry. There are numerous natural products and medicines in their compositions that display such heterocyclic moiety. They are commonly considered as a prototype molecule to many drugs [1–4]. In addition to the pharmaceutical significance, quinoline derivatives have been extensively investigated in different fields, with a wide range of practical applications. Because of their chemical reactivity and thermal stability, electron transfer and emission capacities, and the probability of chemical variations, these compounds are often used in optoelectronics. They are essential for the manufacturing of new OLED devices [5, 6]. Over the past few decades, photochemical and photophysical research has been carried out on aminoquinolines since they evince a broad range of chemical and biological behavior [7–14].

The advancement of nonlinear optical (NLO) materials has gained great prominence in various technologies, including optical computing, dynamic image processing, optical switches, dye lasers, telecommunication devices, and optical disks. Therefore, efforts are underway in modeling various NLO-phores, such as inorganic and organic semiconductors, natural/fabricated nanomaterials, molecular dyes, and polymers [5, 15–19].

The quantum chemical calculations give information concerning the molecular structures, orbital interactions, electronic and vibrational properties. Specifically, the DFT (density functional theory) and its time extension (TD-DFT) has become an important and extensively employed method in understanding structural modifications, reactivity, stability, variations in optical properties via dipole moment changes caused by charge transfer (CT) processes. It imparts a close

✉ Mohan Singh Mehata
msmehata@gmail.com

[1] Photophysics Laboratory, Department of Physics, Centre of Advanced Study, DSB Campus, Kumaun University, Nainital 263002, Uttarakhand, India

[2] Laser Spectroscopy Laboratory, Department of Applied Physics, Delhi Technological University, Delhi-110042, India

[3] Physics Department, Lajpat Rai College, Sahibabad, Ghaziabad 201005, U.P, India

Published online: 24 August 2021

correlation in the theoretical and experimental results and helps to prophesy the molecular properties with accuracy [20–29]. Also, DFT has been considered the simplest and cost-effective method for foretelling the NLO properties of organic materials. It helps to predict the molecular polarizability, hyperpolarizability, and dipole moments of the molecular system [30–36]. The theory, mentioned above, has the advantage of low computational cost, appropriate programs, and excellent results with a range of hybrid functionals. Many theoretical studies, including DFT/TD-DFT-based approaches, were reported [37–48].

In this work, the DFT/TD-DFT methods were employed to 6-aminoquinoline (6AQ), taking into account the significance of quantum chemical studies to analyze different properties of the organic molecule. So far, no computational analysis of 6AQ using the DFT/TDDFT method has been reported. The theories mentioned above were applied to get complete structural features, including optimized structure, geometrical parameters, electrostatic potential (ESP) surface, frontier molecular orbitals, NBO study, various reactivity, and thermodynamic parameters. To explain the CT within the 6AQ molecule, HOMO–LUMO analyses and absorption and emission transitions were investigated. The study has been performed in different solvents and the gas phase. The theoretically obtained results (absorption and emission wavelengths) were correlated with the experimental results [7]. Moreover, the nonlinear optical (NLO) properties, including first hyper-polarizability, polarizability, and dipole moment for 6AQ, were computed using four different hybrid functionals (B3LYP, CAM-B3LYP, PBEPBE, and WB97XZ) to acquire a profound understanding of the relationship between a nonlinear response and molecular framework.

Computational Methodology

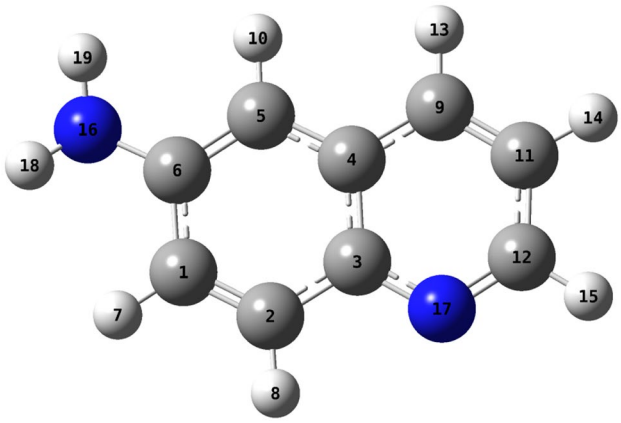
The 6-aminoquinoline (6AQ) was optimized employing the density functional theory (DFT) augmented with B3LYP functional at 6-311G(d,p) with the Gaussian-9 program's package [49]. The molecular structure was drawn using Gauss-view 5.0. To improve the presentation of molecular electron density, the inclusion of these polarization functions was necessary. The optimization energies, geometrical parameters, zero-point corrected energy, and some thermodynamic properties have been estimated to predict the stability of the molecule.

Furthermore, other useful information can also be obtained by the DFT/TD-DFT theory based on the optimized structure, such as interpretation of electrostatic potential (ESP) map, frontier molecular orbitals (FMOs), natural bonding orbital (NBO), nonlinear optical (NLO) and reactivity parameters. Four distinct hybrid functionals, including the long-ranged functionals (B3LYP, CAM-B3LYP, PBEPBE, and WB97XZ), were considered while studying the NLO properties (polarizability and first-order hyper-polarizability) for 6AQ. The study was performed in the gas phase and various solvents (water, ethanol, acetonitrile, chloroform, tetrahydrofuran, benzene, and toluene). Also, the TD-DFT calculations were performed at B3LYP/6-311G(d,p) to describe the electronic transitions. The solvent effect has been considered using an integral equation formalism version of the polarizable continuum model (IEF-PCM).

Results and Discussion

Geometrical Optimization and Thermodynamic Properties

The ground state (GS) (S_0) and the lowest excited state (ES) (S_1) optimization of 6-aminoquinoline (6AQ) was performed using the DFT/TD-DFT approach with 6-311G(d p) basis set and B3LYP functional in the gas phase. Scheme1 represents the molecular structure of 6AQ. The optimized geometric forms corresponding to the local minima without imaginary frequency were confirmed using frequency calculations. According to the structural information, the various calculated C–C and C–H bond lengths between the ring C-atoms as well as the different bond angles, are within the range of the experimental findings for 8-aminoquinoline [50]. The observed bond angles and bond lengths are also well matched with the theoretical results of 5AQ [43]. The geometrical parameters (bond lengths and bond angles) of the 6AQ in the GS (S_0) and the first ES (S_1) are presented in Table S1 of supplementary information (SI). The estimated C–C and C–H bond length range from 1.36 Å to 1.43 Å and 1.08 Å to 1.09 Å respectively. The bond length between C(3) – N(17) and C(12) – N(17) is around 1.36 Å and 1.32



Scheme 1 Molecular structure of 6-aminoquinoline

Å respectively. Considering the different bond angles, the computed angles are very close to each other except near the amino group substitution. The ring's symmetry can be affected by the electron-donating moieties on the benzene ring, resulting in the ring angles less than 120° at the substitution point and marginally greater than 120° at the meta and ortho positions [43]. A similar trend is noticed for 6AQ where the bond angle C(1)–C(6)–C(5) is 118.89° while C(2)–C(1)–C(6) and C(1)–C(2)–C(3) are 121.16° and 120.99°. The bond angles involving the exocyclic and the ring nitrogen atom C(4)–C(3)–N(17), C(1)–C(6)–N(16) and C(5)–C(6)–N(16) are noticed at 123.04°, 119.04° and 121.99°. The inclusion of nitrogen atoms creates major disruptions in the structure of quinoline moiety, while the geometry of the benzene ring remains uninterrupted [43].

However, in the ES, the various bond lengths and angles changed significantly due to the redistribution of charge. For example, the C(1)–C(2) bond length increased from 1.37 Å to 1.39 Å, C(5)–C(6) changes from 1.38 Å to 1.42 Å, C(6)–N(16) decreased from 1.39 Å to 1.36 Å, C(11)–C(12) reduced from 1.42 Å to 1.39 Å. The largest variations are observed in the bond angles near the amino group region. The C(2)–C(1)–C(6) bond angle decreased to 117.07° from 121.16°, C(6)–N(16)–H(18) increased from 115.62° to 120.92°.

The energies and thermodynamic parameters of the titled molecule estimated using frequency calculations by employing B3LYP/6-311G(d,p) at 298 K in the GS are displayed in Table 1. The value of the total dipole moment computed by the given Eq. 1 is higher than quinoline (2.02 D) [51]. This may be attributed due to the existence of the amino group.

Table 1 The thermodynamic properties associated with 6AQ estimated using B3LYP/6311(d,p)

Thermodynamic Parameters (298 K)	6AQ
SCF energy (kcal mol^{-1})	–287,022.66
Total energy, thermal (E_{total}) (kcal mol^{-1})	100.26
Zero point vibrational energy (E_0) (kcal mol^{-1})	95.15
Entropy (S) (cal mol^{-1} K^{-1})	88.09
Specific heat (C_v) (cal mol^{-1} K^{-1})	33.99
Vibrational energy (E_{vib}) (kcal mol^{-1})	98.50
Rotational constants (GHz)	
A	2.89
B	0.84
C	0.65
Dipole moment (Debye)	
μ_x	2.39
μ_y	–2.16
μ_z	0.95
μ_{total}	3.36

$$\mu_{total} = (\mu_x^2 + \mu_y^2 + \mu_z^2)^{1/2} \quad (1)$$

Frontier Molecular Orbitals (MOs) and Global Reactivity Parameters

The FMOs have a meaningful role in forecasting the chemical and kinetic stability, the electrical and optical properties and electronic spectra of the molecules. The HOMO and LUMO signify the tendency to accept and donate an electron. The LUMO reflects the electron-deficient portion of the molecule, while the HOMO reflects the electron-rich portion. Moreover, a lower energy gap (HOMO–LUMO) can explain the possible CT interaction in the molecule [32–35]. The 6AQ molecule has 246 MOs, of which 38 are occupied and 208 are unoccupied. The HOMO/LUMO and band-gap energies estimated for 6AQ are shown in Table 2. The FMOs with an energy difference are displayed in Fig. 1. It can be seen that the HOMO and the LUMO are localized on the entire molecule. This suggests a strong orbital overlap between HOMO/LUMO, facilitating the transition from GS and ES. The calculated HOMO/LUMO energies and their energy gap are −5.68 eV, −1.39 eV and 4.29 eV, respectively.

The HOMO/LUMO energies are considered in calculating the reactivity parameters following the Koopman's theorem [52, 53]. DFT-based global reactivity parameters, for example, ionization potential (IP), chemical potential (μ), electron affinity (EA), electronegativity (χ), electrophilicity index (ω), electron-donating (ω^-) and electron-accepting (ω^+) power and hardness (η) are crucial in exploring the chemical reactivity of molecules in different surroundings and comprehending certain features associated with the reactions.

Table 2 Estimated reactivity parameters including LUMO and HOMO energy and the energy gap using DFT/B3LYP/6-311G(d,p)

Parameters (eV)	6AQ
HOMO	–5.68
LUMO	–1.39
Energy gap (ΔE)	4.29
Ionization Potential (IP)	5.68
Electron Affinity (EA)	1.39
Electronegativity (χ)	3.53
Chemical Potential (μ)	–3.53
Hardness(η)	2.14
Electrophillicity (ω)	2.19
Electron accepting power (ω^+)	1.41
Electron donating power (ω^-)	4.95
Softness (σ) (eV^{-1})	0.23

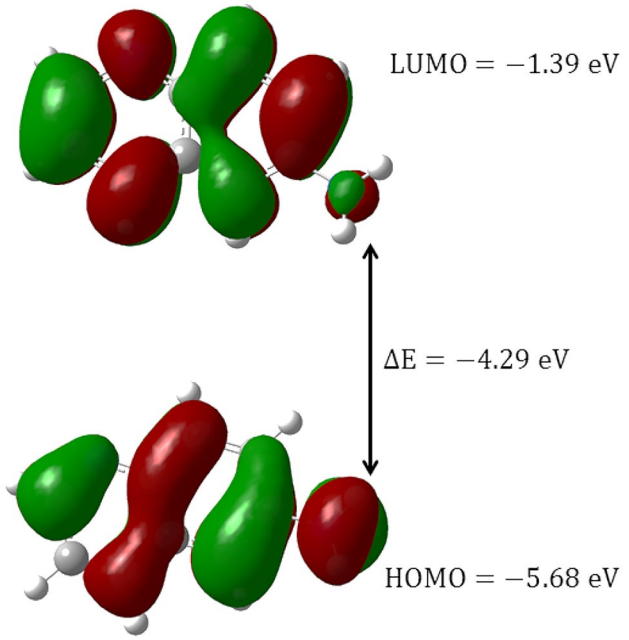


Fig. 1 The HOMO/LUMO of 6AQ using B3LYP/6311-G(d,p)

The following equations are used in defining the reactivity parameters:

$$IP\ (eV) = -E_{HOMO} \tag{2}$$

$$EA(eV) = -E_{LUMO} \tag{3}$$

$$\eta(eV) = \frac{IP - EA}{2} \tag{4}$$

$$\chi(eV) = \frac{IP + EA}{2} \tag{5}$$

$$\mu(eV) = -\chi \tag{6}$$

The electrophilicity index (ω) defined by Parr et al. [54], and electron donating (ω^-) and electron-accepting (ω^+) powers [55] are given by the following Eqs.:

$$\omega(eV) = {}^{\mu^2}/_{2\eta} \tag{7}$$

$$\omega^+(eV) = \frac{(IP + 3EA)^2}{16(IP - EA)} \tag{8}$$

$$\omega^-(eV) = \frac{(3IP + EA)^2}{16(IP - EA)} \tag{9}$$

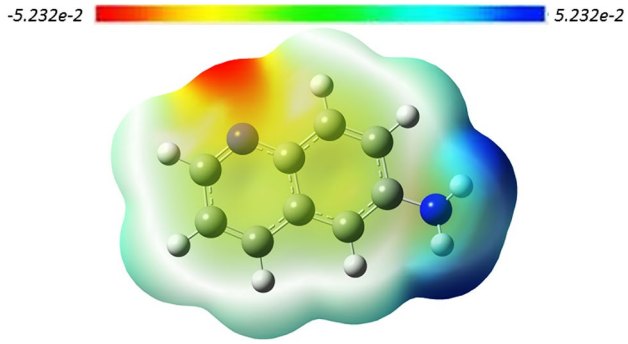


Fig. 2 ESP map for 6AQ computed using B3LYP/6-311G(d,p)

$$\sigma(eV^{-1}) = {}^{1}/_{2\eta} \tag{10}$$

The conceptual global reactivity parameters for the optimized geometry of 6AQ are summarized in Table 2.

The EA and IP are analogous to the electron-accepting and donating property of molecules and are determined by LUMO and HOMO values. The IP value describes the electron-donating ability, indicating that 6AQ has a good donating ability with IP = 5.68 eV. The positive (EA) value is a good indication of the potential use of 6AQ in CT reactions. The χ value relates to the tendency of molecules to attract electrons which is computed to be 3.53 eV. The negative chemical potential (μ = −3.53 eV) shows that the titled molecule is stable. The σ and η are related to the reactivity and the stability of the compound. (ω) represents the molecule's stabilization energy. The molecules with a greater value of hardness are more stable than the molecules with greater softness and are reactive. A molecule with a small value of (ω) is treated as a strong nucleophile, while the molecule with a high value of (ω) is considered as a strong electrophile. For the titled compound, η = 2.14 eV, ω = 2.91 eV and σ = 0.23 eV. The values of (ω^+) and (ω^-) divulges that 6AQ has a high electron-donating nature. According to the electrophilicity index, molecules are categorized as potent electrophiles with ω > 1.5 eV, moderate with 0.8 < ω < 1.5 eV and minor with ω < 0.8 eV [54].

Electrostatic Potential (ESP) Map:

The ESP map is highly informative in reflecting the electrophilic and nucleophilic reactivity regions of a molecule. It is represented by distinct colors where blue signifies the positive (most electron-poor) part, and green indicates the neutral portion and red signifies the negative (most electron-rich) part [33–35]. Figure 2 demonstrates the ESP map of 6AQ and is within the range −5.23e − 2 a.u. to +5.23e − 2 a.u. It can be seen that 6AQ exhibits a strong negative potential (red color) on the ring nitrogen atom and weak negative

potential (yellow color) on the aromatic ring system supporting its contribution as the H-bond acceptor as well as *n* and π donor. The blue region associated with the deficiency of electrons (i.e., nucleophilic reactivity) is around the H-atoms of NH_2 group, a suitable site for inter-molecular H-bonding (nucleophile attack). The ESP map indicates that the electron density (red region) is centralized on the ring nitrogen atom for 6AQ. The nitrogen atom of the quinoline ring has the highest negative potential, making it a favored site for an electrophilic attack.

Electronic Absorption and Emission Spectra

The TD-DFT calculations utilizing B3LYP functional at 6-311G(d,p) basis set were employed to examine the absorption and emission spectra of 6AQ. The experimental absorption and emission wavelengths for 6AQ in few solvents have been reported earlier [7]. The selected solvents in the present calculation are water, ethanol (EtOH), acetonitrile (ACN), tetrahydrofuran (THF), chloroform ($CHCl_3$), benzene, and toluene. The computed vertical excitation energy with the oscillator strength (OS), compositions, and assignment for 6AQ are summarized in Table 3. The HOMO/LUMO have π and π^* nature, respectively. It is to be mentioned that no significant changes were noticed in the delocalization of HOMO and LUMO in different solvents. The main electronic transition band in the UV–Vis spectra of 6AQ results from HOMO → LUMO (with 92–95% contribution) computed between 327 nm–340 nm depending on the medium. This transition between HOMO → LUMO can be classified as $\pi \rightarrow \pi^*$. Figure 3 displays the simulated UV–Vis spectra of 6AQ in the selected solvents as well as the gas phase. The strongest absorption band with λ_{max} leading to $S_0 \rightarrow S_1$ transition is at 327 nm in the gas phase, 335 nm in toluene, 335 nm in benzene, 337 nm in $CHCl_3$, 338 nm in THF, 339 nm in ACN, 340 nm in EtOH and 340 nm in water. The absorption maximum is shifted towards red on increasing the polarity of solvents (Table 3). Calculations were also performed with different hybrid functionals, which reveals that DFT-B3LYP provides more accurate results.

For the emission spectra, the optimization of the lowest ES (S_1) of 6AQ in the gas phase and solvents was performed using the same (TD-DFT/B3LYP/6-311G(d,p)) approach. The PCM model has been used to explore the impact of solvents on the spectral properties of 6AQ. The obtained results are listed in Table S2 of SI. Depending on the medium, the emission wavelength of 6AQ was obtained between 389 nm – 407 nm and ascribed to

Table 3 Excited-state properties, including major transitions and oscillator strength, computed using TD-DFT/B3LYP/6-311G(d,p) for 6AQ

Medium	Electronic transitions	Energy (eV)	λ_{calc}(nm)	OS	Contribution
	$S_0 \rightarrow S_1$	3.77	327	0.0682	H → L (92%)
Gas	$S_0 \rightarrow S_2$	4.29	289	0.0016	H – 2 → L (98%)
	$S_0 \rightarrow S_3$	4.49	276	0.0014	H – 1 → L(55%)
	$S_0 \rightarrow S_1$	3.64	340	0.0925	H → L (95%)
Water	$S_0 \rightarrow S_2$	4.43	279	0.0019	H – 1 → L (53%)
	$S_0 \rightarrow S_3$	4.44	279	0.0019	H – 2 → L(98%)
	$S_0 \rightarrow S_1$	3.65	340	0.0943	H → L (95%)
EtOH	$S_0 \rightarrow S_2$	4.43	279	0.0019	H – 2 → L (98%)
	$S_0 \rightarrow S_3$	4.43	279	0.0019	H – 1 → L(53%)
	$S_0 \rightarrow S_1$	3.65	339	0.0932	H → L (95%)
ACN	$S_0 \rightarrow S_2$	4.29	289	0.0016	H – 1 → L (53%)
	$S_0 \rightarrow S_3$	4.49	276	0.0014	H – 2 → L(97%)
	$S_0 \rightarrow S_1$	3.66	338	0.0973	H → L (95%)
THF	$S_0 \rightarrow S_2$	4.41	281	0.0019	H – 2 → L (98%)
	$S_0 \rightarrow S_3$	4.44	279	0.0017	H – 1 → L(54%)
	$S_0 \rightarrow S_1$	3.67	337	0.0998	H → L (95%)
$CHCl_3$	$S_0 \rightarrow S_2$	4.39	282	0.0019	H – 2 → L (98%)
	$S_0 \rightarrow S_3$	4.45	278	0.0016	H – 1 → L(54%)
	$S_0 \rightarrow S_1$	3.70	335	0.1033	H → L (95%)
Benzene	$S_0 \rightarrow S_2$	4.35	284	0.0019	H – 2 → L (98%)
	$S_0 \rightarrow S_3$	4.46	278	0.0015	H – 1 → L(55%)
	$S_0 \rightarrow S_1$	3.69	335	0.1030	H → L (94%)
Toluene	$S_0 \rightarrow S_2$	4.35	284	0.0019	H – 2 → L (98%)
	$S_0 \rightarrow S_3$	4.46	278	0.0015	H – 1 → L(55%)

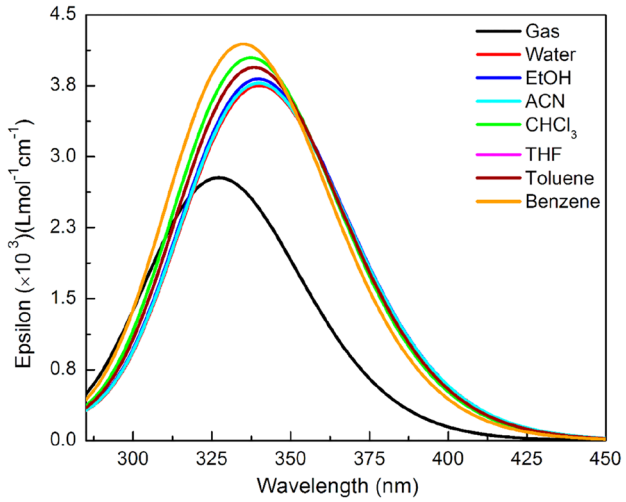


Fig. 3 TD-DFT-based UV–vis spectra of 6AQ in the selected solvents as well as in the gas phase at DFT/B3LYP/6-311G(d,p) level

$S_1 \rightarrow S_0$ transition. Following the pattern of absorption spectra, $S_1 \rightarrow S_0$ transitions in emission spectra have the highest oscillator strength and emerge mainly from LUMO → HOMO transition. It is observed that with increasing solvent polarity, the emission maximum shows a continuous increase in the wavelength. Thus it is observed that the lowest ES is $\pi \rightarrow \pi^*$, CT state, which is consistent with the literature [7]. The bathochromic shift observed experimentally is in harmony with the present theoretical results. Fig. 4 and S1 (SI) show the simulated emission spectra in solvents as well as the gas phase. The S_0 and S_1 state dipole moment in different solvents and the gas phase are given in Table S3 of SI. The theoretically calculated dipole moment for 6AQ in the ES is higher than in the GS. This implies that 6AQ is more polar in its ES state than in the GS state and that the ES state is more stable than the GS state in polar solvents, compared to aprotic and non-polar solvents.

The absorption and emission spectra of 6AQ in the selected solvents were recorded experimentally. The results are well-matched with the reported values of EtOH, ACN, $CHCl_3$ and water at specific pH [7]. The obtained experimental and theoretical results and the reported values are summarized in Table S4 of SI.

The computed absorption wavelength of 6AQ in water is well-matched with the experimental results; however, a variation of 10–15 nm from the experimental data is observed for other studied solvents. Emission wavelength in ACN and $CHCl_3$ is well-matched with the experimental results though large variation can be seen in the emission wavelength for water and EtOH. The difference in theoretical and experimental values can be attributed to various assumptions in theoretical approaches. The theoretically computed Stokes shift in the specified solvents is closely related to the solvent polarity function shown in Table S5 of SI. The analysis shows a linear association between $(\overline{\nu}_a - \overline{\nu}_f)$ and $F(\varepsilon, \eta)$ in the chosen solvents (Fig. 5). Generally, the larger the $F(\varepsilon, \eta)$, the larger the Stokes shift. For this study, the following (Eq. 11) Lippert-Mataga polarity function was utilized [57, 58].

$$F(\varepsilon, \eta) = \left[\frac{\varepsilon - 1}{2\varepsilon - 1} - \frac{\eta^2 - 1}{2\eta^2 - 1}\right] \tag{11}$$

The accumulation of the charge on the ring nitrogen atom in S_1 state reinforces the H-bonding of the lone pairs

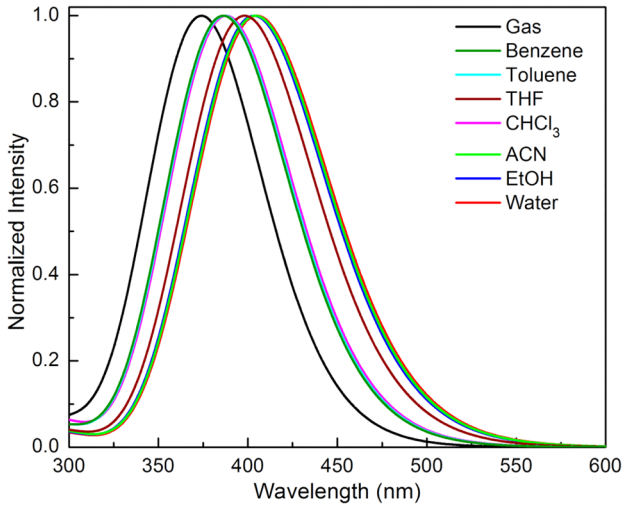


Fig. 4 Normalized emission spectra of 6AQ estimated using TD-DFT/B3LYP/6-311G(d,p) in selected the solvents and gas phase

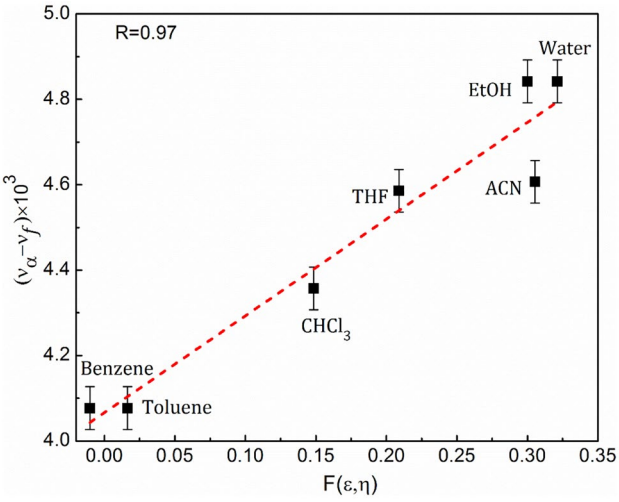


Fig. 5 TD-DFT-based correlation between Stoke's shift and solvent polarity function $\boldsymbol{F(\varepsilon, \eta)}$

Table 4 Second-order perturbation energies $E^{(2)}$ (kcal/mol) of few donor–acceptor interactions in 6AQ

Donor(i)	Acceptor(j)	$E^{(2)}$	$\varepsilon_j - \varepsilon_i$	F_{ij}
BD(2) C(1)-C(2)	BD ∗ (2) C(3) – C(4)	14.44	0.29	0.062
BD(2) C(1)-C(2)	BD ∗ (2) C(5) – C(6)	18.85	0.30	0.069
BD(2) C(3)-C(4)	BD ∗ (2) C(1) – C(2)	16.85	0.28	0.065
BD(2) C(3)-C(4)	BD ∗ (2) C(5) – C(6)	15.53	0.27	0.060
BD(2) C(3)-C(4)	BD ∗ (2) C(9) – C(11)	17.58	0.27	0.066
BD(2) C(3)-C(4)	BD ∗ (2) C(12) – N(17)	16.11	0.26	0.060
BD(2) C(5)-C(6)	BD ∗ (2) C(1) – C(2)	14.03	0.30	0.058
BD(2) C(5)-C(6)	BD ∗ (2) C(3) – C(4)	19.65	0.29	0.071
BD(2) C(9)- C(11)	BD ∗ (2) C(3) – C(4)	15.07	0.30	0.063
BD(2) C(9)-C (11)	BD ∗ (2) C(12) – N(17)	22.55	0.28	0.072
BD(2) C(12)-N(17)	BD ∗ (2) C(3) – C(4)	19.03	0.33	0.076
BD(2) C(12)-N(17)	BD ∗ (2) C(9) – C(11)	12.38	0.33	0.058
LP(1) N(16)	BD ∗ (2) C(5) – C(6)	28.82	0.33	0.092
LP(1) N(17)	BD ∗ (2) C(3) – C(4)	10.63	0.86	0.086
LP(1) N(17)	BD ∗ (2) C(11) – C(12)	9.94	0.87	0.084

and promotes the redshift shift of the emission on increasing the polarity. It also demonstrates that the H-bond to the lone pairs of the amino nitrogen atom in S_1 state, in the solvent cage, is smashed by the relocation of lone pairs to the ring upon excitation [7]. Thus, the allowed $\pi \rightarrow \pi^*$ transition shows the CT property.

Natural Bonding Orbital Analysis

Natural bonding orbital (NBO) study provides an appropriate framework for investigating the transfer of charge and intra- or inter-molecular bonding in molecular systems. The stable interaction between donor and acceptor involves the electron density delocalization of occupied and unoccupied NBOs [5, 13, 41, 56]. The NBO analysis of 6AQ at DFT/B3LYP/6-311G(d,p) was performed to interpret the CT from donor to acceptor. The stabilization energy of a molecule in NBO study is calculated using second-order perturbation theory and is defined by the given equation:

$$E^{(2)} = q_i \frac{(F_{ij})^2}{\varepsilon_j - \varepsilon_i} \tag{12}$$

where, $E^{(2)}$ signifies stabilization energy, q_i symbolizes orbital occupancy of the donor (ε_i,ε_j) and F_{ij} is the off-diagonal and diagonal NBO Fock matrix elements [5, 41]. The more the value of $E^{(2)}$, the stronger is the association between electron acceptor and donor moieties [41]. From Table 4, it is noticed that the CT interaction of lone pairs of exocyclic nitrogen atom LP(1)N(16) with the antibonding (π^*) C(5)–C(6) has higher stabilization energy. The large energy value is due to a small gap in the energy of the electron donor and acceptor. Thus, the NBO investigation of the titled molecule reveals that the ICT is significant in stabilizing the molecular system.

Nonlinear Optical (NLO) Properties:

To comprehend the nonlinear behavior, we computed the polarizability (α) and first-order hyperpolarizability (β) of 6AQ at B3LYP, Cam-B3LYP, WBP7XB, and PBEPBE functional using 6-311G(d,p). The computed values are tabulated in Tables 5 and 6. The polarizability (α), first-order hyperpolarizability (β), is determined by equations below, using the x, y, z components [5, 32]:

$$\alpha_{total} = \frac{1}{3}(\alpha_{xx} + \alpha_{yy} + \alpha_{zz}) \tag{13}$$

$$\beta_{\text{total}} = (\beta_x^2 + \beta_y^2 + \beta_z^2)^{1/2} \tag{14}$$

where,

$$\beta_x = \beta_{xxx} + \beta_{xyy} + \beta_{xzz}$$

$$\beta_y = \beta_{yyy} + \beta_{yzz} + \beta_{yxx}$$

Table 5 The polarizability (α) for 6AQ calculated using selected functionals and 6-311G(d,p) basis set

Polarizability(α)	B3LYP	CAM – B3LYP	WB97XB	PBEPBE
α_{xx}(a.u.)	182.60	174.86	175.08	192.11
α_{yy}(a.u.)	114.53	111.67	112.24	118.60
α_{zz}(a.u.)	46.26	45.92	46.19	46.89
α_{xy}(a.u.)	5.43	5.24	5.211	5.34
α_{xz}(a.u.)	0.30	0.27	0.22	0.26
α_{yz}(a.u.)	−0.018	−0.001	−0.007	−0.034
α_{total}(a.u.)	114.46	110.82	111.17	119.20
α_{total}(e.s.u)	16.96×10^{-24}	16.42×10^{-24}	16.47×10^{-24}	17.66×10^{-24}

Table 6 First-order hyperpolarizability (**β**) at 6-311G(d,p) using different functionals for 6AQ

First-order hyperpolarizability (**β**)	B3LYP	CAM-B3LYP	WB97XB	PBEPBE
β_{xxx}(a.u.)	−744.39	−663.66	−650.45	−824.28
β_{yyy}(a.u.)	−33.29	−24.234	−24.37	−48.01
β_{zzz}(a.u.)	5.92	5.18	5.92	5.14
β_{xxy}(a.u.)	48.96	50.34	48.17	39.58
β_{xyy}(a.u.)	20.32	18.11	17.29	29.82
β_{xxz}(a.u.)	45.99	37.64	38.24	50.89
β_{xzz}(a.u.)	−0.55	−0.28	−0.96	2.42
β_{yzz}(a.u.)	−6.83	−6.07	−5.78	−5.76
β_{yyz}(a.u.)	6.89	5.31	4.69	6.94
β_{xyz}(a.u.)	−4.54	−0.28	−4.27	−6.37
β_{total}(a.u.)	727.06	647.22	636.25	794.86
β_{total}(e.s.u)	6.28×10^{-30}	5.59×10^{-30}	5.49×10^{-30}	6.87×10^{-30}

$$\beta_z = \beta_{zzz} + \beta_{zxx} + \beta_{zyy}$$

therefore, Eq. 14. is written as:

$$\beta_{\text{total}} = \left[\left(\beta_{xxx} + \beta_{xyy} + \beta_{xzz}\right)^2 + \left(\beta_{yyy} + \beta_{yzz} + \beta_{yxx}\right)^2 + \left(\beta_{zzz} + \beta_{zxx} + \beta_{zyy}\right)^2 \right]^{1/2}$$

The first-order hyperpolarizability is represented as a $3 \times 3 \times 3$ matrix and is a tensor of rank 3. According to Kleinman [59] symmetry, the twenty-seven components of the $3 \times 3 \times 3$ matrix can be lessened to ten components. From the frequency calculations, the 10 elements are: β_{xxx}, β_{xxy}, β_{xyy}, β_{yyy}, β_{xxz}, β_{xyz}, β_{yyz}, β_{xzz}, β_{yzz}, β_{zzz}, respectively.

The theoretically estimated value of α_{total} and β_{total} at DFT/B3LYP/6-311G(d,p) level is 16.96×10^{-24} e.s.u. and 6.28×10^{-30} e.s.u. The β_{total} value for 6AQ is found to be significantly higher than urea $\left(3.71 \times 10^{-31}\right)$ e.s.u. Though urea is considered a reference molecule in the analysis of NLO properties and, thus, is commonly seen as a standard value for comparisons [35, 60, 61]. The α_{total} value reported for 3-aminoquinoline (18.72×10^{-24} e.s.u.) computed using DFT/B3LYP/6–311 + + G(d.p) method is nearer to the present value for 6AQ [41]. The large hyperpolarizability value is directly correlated to the charge transfer (ICT) properties, emanating from the transition of electron cloud from a region of electron donor to the region of electron acceptor [33, 34].

Molecular geometry and the surrounding environment can influence the NLO properties. Besides, the polarizability and first-order hyperpolarizability of 6AQ determined by DFT at B3LYP/6–311(d,p) in water, EtOH, $CHCl_3$, and toluene are given in Table S6 of SI. The values of dipole moment (Table S3 of SI), α_{total} and β_{total} increases with increasing the solvent polarity. A significant variation is noticed among these values in the gas phase and the preferred solvents. It can be inferred that the ICT character increases as the β_{total} value increases. Therefore, the NLO response of 6AQ can be superficially tuned by varying the solvent polarity.

The β_{total} values for 8-aminoquinoline, 8-hydroxyquinoline, and 8-hydroxyquinolidine determined experimentally using the solvatochromic methods were found to be 10.0×10^{-30} e.s.u., 2.24×10^{-30} e.s.u, and 5.29×10^{-30} e.s.u. [62]. These experimental results of quinoline derivatives can be compared with the present theoretical results for 6AQ. The β_{total} obtained for 6AQ matches well with the experimental value of 8AQ in non-polar solvents, while it is higher in protic and aprotic solvents. However, the observed values of β_{total} for 8HQ and 8HQD were less than 6AQ in solvents. Also, the theoretical value of β_{total} reported for 8HQ [63] and 1-chloroisoquinoline [39] in the gas phase is smaller than the titled molecule. Thus, the estimated values using the methods mentioned above indicate that 6AQ can be considered a desirable molecule for further studies of the NLO properties.

Conclusion

Using DFT/TD-DFT method at B3LYP/6-311G(d,p) level, the optimized structure of 6AQ in both the GS and ES along with electronic spectra, structural parameters, electrostatic potential (ESP) surface, molecular orbitals (MOs), various reactivity and thermodynamic parameters were estimated. The ESP surface of the titled molecule shows that the largest negative potential exists around the quinoline's ring nitrogen atom and is considered the favored site of the electrophilic attack. The influence of various solvents in the absorption and emission spectra of 6AQ using the PCM model was analyzed and correlated with the experimental findings. Depending on the medium, the computed absorption maxima of 6AQ appear between 327 nm—340 nm and ascribed to $S_0 \rightarrow S_1$ transitions. The simulated emission maxima were computed between 376 nm – 410 nm and ascribed to $S_1 \rightarrow S_0$ transitions. The lowest ES is noticed as $\pi \rightarrow \pi^*$ CT state. The natural bonding orbital (NBO) study revealed that ICT contributes significantly in stabilizing the molecular system. Furthermore, the NLO properties (polarizability, dipole moment and first-order hyperpolarizability) estimated using different hybrid functionals suggest that 6AQ can be considered in NLO applications.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10895-021-02788-z.

Acknowledgements Author NP is grateful to the University Grant Commission (UGC), New Delhi, India, for supporting under the RFSMS (Research Fellowship in Science for Meritorious students) scheme (Sanction no. F. No. 25-1/2014-15(BSR)5–53/2007). Author MSM is thankful to Prof. Keli Han, Dalian Institute of Chemical Physics, China, for the Gaussian 09 facilities.

Authors Contributions NP: Conception and design the study, acquisition of data and drafting the manuscript. MSM: Resources, funding acquisition, supervision, revising and editing the manuscript. SP: Supervision, read the final version of the manuscript and provided valuable discussions. NT: Revising and editing the manuscript.

Data Availability All data generated or analysed during this study are included in this published article and its supplementary information file.

Code Availability Gaussian 09 W and Gauss view 5 are used for simulation and origin software is used to draw the plots.

Declarations

Ethical Approval This article does not contain any studies with human participants or animals, clinical trial registration or plant reproducibility performed by any authors.

Consent to Publish All authors have approved the paper and agree with its publication.

Conflict of Interest The authors declare that they have no conflicts of interest.

References

1. Dos Santos GC, Servilha RO, de Oliveira EF et al (2017) Theoretical-experimental photophysical investigations of the solvent effect on the properties of green-and blue-light-emitting quinoline derivatives. J Fluoresc 27:1709–1720
2. Joshi NK, Joshi HC, Gahlaut R et al (2012) Steady state and time-resolved fluorescence study of isoquinoline: reinvestigation of excited state proton transfer. J Phys Chem A 116:7272–7278
3. Marella A, Tanwar OP, Saha R et al (2013) Quinoline: A versatile heterocyclic. Saudi Pharm J 21:1–12
4. Arjunan V, Balamourougane PS, Kalaivani M et al (2012) Experimental and theoretical quantum chemical investigations of 8-hydroxy-5-nitroquinoline. Spectrochim Acta Part A Mol Biomol Spectrosc 96:506–516
5. Khalid M, Ali A, Jawaria R et al (2020) First principles study of electronic and nonlinear optical properties of A-D–π–A and D–A–D–π–A configured compounds containing novel quinoline–carbazole derivatives. RSC Adv 10:22273–22283
6. Danel A, Gondek E, Kityk I (2009) 1H-pyrazolo [3, 4-b] quinoline and 1H-pyrazolo [3, 4-b] quinoxaline derivatives as promising materials for optoelectronic applications. Opt Mater (Amst) 32:267–273
7. Schulman SG, Abate K, Kovi PJ et al (1973) Photoluminescence of 6-and 7-Aminoquinolines. Anal Chim Acta 65:59–67
8. Schulman SG, Sanders LB (1971) Fluorescence and phosphorescence of 5- and 8-minoquinoline. Anal Chim Acta 56:83–89
9. Pandey N, Mehata MS, Fatma N, Pant S (2019) Efficient fluorescence quenching of 5-aminoquinoline: Silver ion recognition study. J Lumin 205:475–481
10. Bridhkoti JP, Joshi HC, Pant S (2012) Polymer micro-environmental effects on the fluorescence characteristics of 5-aminoquinoline and 3-aminoquinoline. J Lumin 132:722–729
11. Das S, Singh AK, Biswas DS, Datta A (2019) Dynamics of Preferential Solvation of 5-Aminoquinoline in Hexane-Alcohol Solvent Mixtures. J Phys Chem B 123:10267–10274
12. Kovi PJ, Capomacchia AC, Schulman SG (1972) Electronic spectra of 2-aminoquinoline and 4-aminoquinaldine. Evidence for the cyclic amidine structures of the singly protonated cations. Anal Chem 44:1611–1615
13. Panda D, Ghosh D, Datta A (2009) Acid-base behavior of 3-aminoquinoline in its ground and excited states. J Photochem Photobiol A Chem 207:254–259
14. Iwata K, Kakita M, Hamaguchi H (2007) Picosecond Time-Resolved Fluorescence Study on Solute− Solvent Interaction of 2-Aminoquinoline in Room-Temperature Ionic Liquids: Aromaticity of Imidazolium-Based Ionic Liquids. J Phys Chem B 111:4914–4919
15. Fonseca RD, Vivas MG, Silva DL et al (2018) First-order hyperpolarizability of triphenylamine derivatives containing cyanopyridine: molecular branching effect. J Phys Chem C 122:1770–1778
16. Halasyamani PS, Zhang W (2017) Inorganic materials for UV and deep-UV nonlinear-optical applications.12077–12085
17. Hussain A, Khan MU, Ibrahim M et al (2020) Structural parameters, electronic, linear and nonlinear optical exploration of thiopyrimidine derivatives: A comparison between DFT/TDDFT and experimental study. J Mol Struct 1201:127183
18. Deb J, Paul D, Sarkar U (2020) Density functional theory investigation of nonlinear optical properties of T-graphene quantum dots. J Phys Chem A 124:1312–1320
19. Guo L, Guo Z, Li X (2018) Design and preparation of side chain electro-optic polymeric materials based on novel organic second order nonlinear optical chromophores with double carboxyl groups. J Mater Sci Mater Electron 29:2577–2584
20. Alyar H (2013) A review on nonlinear optical properties of donor-acceptor derivatives of naphthalene and azanaphthalene. Rev Adv Mater Sci 34:e87
21. Eşme A (2019) Experimental (FT-IR, FT-Raman, and UV-Vis) and quantum chemical calculations on monomer and dimer structures of l-hydroxy-2-naphthoic acid using the DFT and TD-DFT methods. Indian J Pure Appl Phys 57:822–835
22. Wazzan NA (2018) Investigation on the performance of PCM/TD-DFT functionals (standard pure, hybrid and long-range corrected) in simulating the absorption and emission spectra of 1-(2, 5-dimethylfuran-3-yl)-3-(2, 4, 5-trimethoxyphenyl) prop-2-en-1-one dye in different solv. J Mol Struct 1164:289–296
23. Liu X, Yang D, Ju H et al (2011) Experimental and PCM/TD-DFT investigation on the absorption and emission spectra of a light emitting material in various solvents. Chem Phys Lett 503:75–79
24. Amat A, Miliani C, Romani A, Fantacci S (2015) DFT/TDDFT investigation on the UV-vis absorption and fluorescence properties of alizarin dye. Phys Chem Chem Phys 17:6374–6382
25. Li G, Han K (2018) The sensing mechanism studies of the fluorescent probes with electronically excited state calculations. Wiley Interdiscip Rev Comput Mol Sci 8:e1351
26. Zhao W, Pan L, Bian W, Wang J (2008) Influence of solvent polarity and hydrogen bonding on the electronic transition of coumarin 120: A TDDFT study. ChemPhysChem 9:1593–1602
27. Örücü H, Acar N (2015) Effects of substituent groups and solvent media on Pyrene in ground and excited states: A DFT and TDDFT study. Comput Theor Chem 1056:11–18

28. Li Y, Sun C, Han J et al (2020) Theoretical investigation of intermolecular hydrogen bond induces fluorescence quenching phenomenon for Coumarin-1. J Lumin 221:117110
29. Zhang J, Peng S, Wei Y, Zheng S (2020) A theoretical study of the absorption spectra of electron-deficient pentacene derivatives using DFT and TDDFT. Spectrochim Acta Part A Mol Biomol Spectrosc 225:117480
30. Lanke SK, Sekar N (2016) Pyrazole based NLOphores: synthesis, photophysical, DFT, TDDFT studies. Dye Pigment 127:116–127
31. Ayare NN, Shukla VK, Sekar N (2020) Charge transfer and nonlinear optical properties of anthraquinone D-π-A dyes in relation with the DFT based molecular descriptors and perturbational potential. Comput Theor Chem 1174:112712
32. Rahmani R, Boukabcha N, Chouaih A et al (2018) On the molecular structure, vibrational spectra, HOMO-LUMO, molecular electrostatic potential, UV–Vis, first order hyperpolarizability, and thermodynamic investigations of 3-(4-chlorophenyl)-1-(1yridine-3-yl) prop-2-en-1-one by quantum chemistry calculati. J Mol Struct 1155:484–495
33. Arivazhagan M, Jeyavijayan S (2011) Vibrational spectroscopic, first-order hyperpolarizability and HOMO, LUMO studies of 1, 2-dichloro-4-nitrobenzene based on Hartree-Fock and DFT calculations. Spectrochim Acta Part A Mol Biomol Spectrosc 79:376–383
34. Arivazhagan M, Thilagavathi G (2012) Vibrational study, first hyperpolarizability and HOMO–LUMO analyses on the structure of 2-hydroxy-6-nitro toluene. Spectrochim Acta Part A Mol Biomol Spectrosc 91:411–418
35. Balachandran V, Nataraj A, Karthick T (2013) Molecular structure, spectroscopic (FT-IR, FT-Raman) studies and first-order molecular hyperpolarizabilities, HOMO–LUMO, NBO analysis of 2-hydroxy-p-toluic acid. Spectrochim Acta Part A Mol Biomol Spectrosc 104:114–129
36. Gümüs HP, Tamer Ö, Avci D et al (2014) Theoretical investigations on nonlinear optical and spectroscopic properties of 6-(3, 3, 4, 4, 4-pentafluoro-2-hydroxy-1-butenyl)-2, 4-pyrimidinedione: An efficient NLO material. Russ J Phys Chem A 88:2348–2358
37. Liu Y-H, Mehata MS, Lan S-C (2014) TDDFT study of the polarity controlled ion-pair separation in an excited-state proton transfer reaction. Spectrochim Acta Part A Mol Biomol Spectrosc 128:280–284
38. Sureshkumar B, Mary YS, Suma S et al (2018) Spectroscopic characterization of 8-hydroxy-5-nitroquinoline and 5-chloro-8-hydroxy quinoline and investigation of its reactive properties by DFT calculations and molecular dynamics simulations. J Mol Struct 1164:525–538
39. Vidhya V, Austine A, Arivazhagan M (2020) Molecular structure, aromaticity, vibrational investigation and dual descriptor for chemical reactivity on 1-chloroisoquinoline using quantum chemical studies. Results Mater 6:100097
40. Sureshkumar B, Mary YS, Resmi KS et al (2018) Spectroscopic characterization of hydroxyquinoline derivatives with bromine and iodine atoms and theoretical investigation by DFT calculations, MD simulations and molecular docking studies. J Mol Struct 1167:95–106
41. Soliman SM, Kassem TS, Badr AMA et al (2014) Quantum chemical study, spectroscopic investigations, NBO and HOMO–LUMO analyses of 3-aminoquinoline (3AQ) and [Ag (3AQ) 2 (TCA)] complex (TCA= Trichloroacetate). J Mol Struct 1074:168–179
42. Arjunan V, Ravindran P, Rani T, Mohan S (2011) FTIR, FT-Raman, FT-NMR, ab initio and DFT electronic structure investigation on 8-chloroquinoline and 8-nitroquinoline. J Mol Struct 988:91–101
43. Arjunan V, Mohan S, Balamourougane PS, Ravindran P (2009) Quantum chemical and spectroscopic investigations of 5-aminoquinoline. Spectrochim Acta Part A Mol Biomol Spectrosc 74:1215–1223
44. Bouzian Y, Karrouchi K, Sert Y et al (2020) Synthesis, spectroscopic characterization, crystal structure, DFT, molecular docking and in vitro antibacterial potential of novel quinoline derivatives. J Mol Struct 1209:127940
45. Mehata MS, Singh AK, Sinha RK (2017) Investigation of charge-separation/change in dipole moment of 7-azaindole: Quantitative measurement using solvatochromic shifts and computational approaches. J Mol Liq 231:39–44
46. Tiwari K, Arora P, Pandey N et al (2014) Experimental and computational approaches on dipole moment of 5-aminoisoquinoline. J Mol Liq 200:460–464
47. Soliman SM (2012) Molecular structure and vibrational spectra of free and coordinated 3-bromoquinoline: Unexpected intramolecular CH··· O interactions. J Mol Struct 1017:135–142
48. Pandey N, Mehata MS, Fatma N, Pant S (2020) Modulation of Fluorescence properties of 5-Aminoquinoline by Ag^{+} in aqueous media via charge transfer. J Photochem Photobiol A Chem 112549
49. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Petersson GA, Nakatsuji H, Li X, Caricato M, Marenich AV, Bloino J, Janesko BG, Gomperts R, Mennucci B, Hratchian HP, Ortiz JV, Izmaylov AF, Sonnenberg JL, Williams-Young D, Ding F, Lipparini F, Egidi F, Goings J, Peng B, Petrone A, Henderson T, Ranasinghe D, Zakrzewski VG, Gao J, Rega N, Zheng G, Liang W, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Throssell K, Montgomery JA, Peralta JE Jr., Ogliaro F, Bearpark MJ, Heyd JJ, Brothers EN, Kudin KN, Staroverov VN, Keith TA, Kobayashi R, Normand J, Raghavachari K, Rendell AP, Burant JC, Iyengar SS, Tomasi J, Cossi M, Millam JM, Klene M, Adamo C, Cammi R, Ochterski JW, Martin RL, Morokuma K, Farkas O, Foresman JB, Fox DJ (2009) Gaussian, Inc., Wallingford CT
50. Kisiel Z, Desyatnyk O, Pszczółkowski L et al (2003) Rotational spectra of quinoline and of isoquinoline: spectroscopic constants and electric dipole moments. J Mol Spectrosc 217:115–122
51. Van Meervelt L, Goethals M, Leroux N, Zeegers-Huyskens T (1997) X-ray and vibrational studies of 8-aminoquinoline. Evidence for a three-center hydrogen bond. J Phys Org Chem 10:680–686
52. Koopmans T (1934) Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den einzelnen Elektronen eines Atoms. Physica 1:104–113
53. Eşme A, Sağdınç SG (2017) Spectroscopic (FT–IR, FT–Raman, UV–Vis) analysis, conformational, HOMO-LUMO, NBO and NLO calculations on monomeric and dimeric structures of 4–pyridazinecarboxylic acid by HF and DFT methods. J Mol Struct 1147:322–334
54. Parr RG, Szentpály L, v, Liu S, (1999) Electrophilicity index. J Am Chem Soc 121:1922–1924
55. Gazquez JL, Cedillo A, Vela A (2007) Electrodonating and electroaccepting powers. J Phys Chem A 111:1966–1970
56. Kavitha E, Sundaraganesan N, Sebastian S, Kurt M (2010) Molecular structure, anharmonic vibrational frequencies and NBO analysis of naphthalene acetic acid by density functional theory calculations. Spectrochim Acta Part A Mol Biomol Spectrosc 77:612–619
57. Lippert E (1957) Spectroscopic determination of the dipole moment of aromatic compounds in the first excited singlet state. Z Elektrochem 61:962–975
58. Mataga N, Kaifu Y, Koizumi M (1956) Solvent effects upon fluorescence spectra and the dipole moments of excited molecules. Bull Chem Soc Jpn 29:465–470

59. Kleinman DA (1962) Nonlinear dielectric polarization in optical media. Phys Rev 126:1977
60. Mishra AK, Tewari SP (2020) Density functional theory calculations of spectral, NLO, reactivity, NBO properties and docking study of Vincosamide-N-Oxide active against lung cancer cell lines H1299. SN Appl Sci 2:1–13
61. Sakthi D, Prakasam M, Prakasam A et al (2017) A Complete DFT, TD-DFT and Non-Linear Optical Property Study on 6-Amino-2-Methylpyridine-3-Carbonitrile. Comput Chem 5:129
62. Purohit G, Joshi GC (2003) Second-order polarizabilities of some quinolines. Indian J Pure Appl Phys 41:922–927
63. Camargo AJ, Napolitano HB, Zukerman-Schpector J (2007) Theoretical investigation of the intramolecular hydrogen bond formation, nonlinear optic properties, and electronic absorption spectra of the 8-hydroxiquinoline. J Mol Struct THEOCHEM 816:145–151

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

Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 04, April-2021

Supply Chain Management Analysis in Online Retailers and Logistics

Siddhant Changkakati, Shubhank Singh Chauhan, Shivam Verma, Dr. N Yuvraj
Mechanical Engineering,
Delhi Technological University

***Abstract*:- The world is changing, with invasion of the novel corona virus there has been a complete shift in humanity. The pattern in which society functions has been completely altered. This has had a great impact in the day to day businesses. This paper that we have researched might help normal businesses find their solutions to adopt with the incoming changes and challenges that came along with the onset of the following decade of 2020s. There is acute trend and a prediction that has been observed that online retailers are thriving these changing times. Sooner or later most businesses have to adapt to these changing times. This paper is going to study the aspect of logistics and supply chains of various online retailers that have succeed in bring a level of success in their business. Tangible and particle solutions will be provided. Different strategies and techniques are studied to bring out best possible solutions are being highlighted in this paper. Through our knowledge in supply chain management that we have gathered in our college education we have tried to give our solutions on various challenges faced by businesses.**

Key Words - Supply Chain Management (SCM), logistics, online retailer

I. INTRODUCTION

Supply chain management is an interesting subject. Physical shopping will be lessening in the coming decade. There is a greater need to make the product and service available to the consumers to their door steps. To make this process a success requires an intricate design of the supply chain. This paper is written in such a way as to guide the reviewer/reader to make a proper conclusion that will help their businesses thrive in online platforms. Before we start we would like to make you aware of the topics of supply chain that will help to progress further with the paper.

A. SCM TOPICS

- Facility Location:- This involves with the strategic placement of various facility locations. There are number factors that determines the locations of various facilities like proximity to the market, business climate, labor etc. the various ways of analysis of optimization of facility location are-
    i. transportation method
    ii. Centroidal method
    iii. Cost-volume-profit ratio/ break- even analysis
    iv. Factor rating method

- Plant Layout:- This involves with arrangement of the facility in order to have the maximum amount of productivity. There are different types of plant layout structures such as process layout, product layout group layout etc. various ways that are used to analyze and optimize are:-
    i. systematic layout planning
    ii. Block diagram
    iii. Assembly line balancing

- Aggregate planning:- It deals with the subject of employment and manufacture and management of the product depending on the on the demand, supply and need of work force. It basically deals with the optimization of the workforce and product management based on the demand and supply.
    i. Work force level in industries ii. Subcontracting
    ii. Overtime
    iii. Hiring and firing

- Inventory control:- To meet the demand in the market certain level of stock of the item is kept. It is classified according to the time it is required to sell of and the availability of the item. Thus it is classified into fast moving, slow moving, non moving and scarce, difficult, easy. There are various EOQ (Economic Order Quantity) models used to calculate the inventory.

- Forecasting:- Inventory management, manufacturing requires an amount of forecasting. This is done using the data of past experiences. There are two methods of forecasting quantitative and qualitative methods.

II. METHODOLOGY

A. Literature review

According to a paper published in iraj journal in 2014 by Ashu Nagpal and Bineet Singha the major challenges faces by retailer in Delhi NCR where highlighted. A survey was conducted on the various problems faced by the retailers. Lack of infrastructure and logistics had a mean rating of 3.91, inefficiency of supply chain had a mean rating of 3.94, and completion from e retailers had a mean ration of 3.49.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 04, April-2021

It is observed that in India the majority of the retailer are unorganized retailers. The retail industry in India is 200 billion US dollar industry. Organized retailers have a share of 3 percent of it which is 6.4 billion US dollars. The major problem faced by retailer in India is mostly related to supply chain and logistics requirement.
According to a paper published in IJRECT by Tanjeet Singh there is a big booming of the ecommerce industry. The facility of the internet has been used by various industries. It expands the market in national and international level. It improves the brand image of the company. Mode of transaction of money becomes easier. It reduces the exposure to outside environment. The customers are able to better review the product. But the main remain obstacle why retailers don't go for e-commerce platform is because they the basic idea and the knowledge and the expertise to venture into such platforms.

More than 60 percent of the workforces in India is in the agricultural sector. .A paper was published in the journal research gate by M. Balakrishnan, B. Ganesh Kumar, Ch. Srinivasa Rao and S.K. Soam. This paper studies the aspects of agriculture and e-commerce. The main problems they face are lack of knowledge and literacy, the lack of electricity, lack of proper supply chain and lack of digital facilities.

One of the oldest and the largest manufacturing sector in India is the textile industry. It is also a vision of the government to digitalize India and to promote 'make in India'. The senerio of textile industries largly consist of handlooms.
The role of ecommerce in the clothing industry according to a paper published in researchgate by Ashok Kumar Panigrahi and Vijay Joshi in 2016 the benefits e-commerce can bring to clothing industry is generate employment, facilitate growth of allied industries, promote entrepreneurship, reduce transaction costs, easy access to consumer market.

B. Market scenario

Types of retailers in India

a) Itinerant retailers
   i. Hawkers and Peddlers
   ii. Market traders
   iii. Street trader
   iv. Cheap jack
b) Fixed small shop retailers
   i. General store
   ii. Specialty store
   iii. Second hand goods shops
   iv. Seconds store
   v. Single line stores
c) Fixed shop large retailer

   i. Departmental stores
   ii. Multiple shops chains stores
   iii. Mail order retailing
   iv. Consumer Co-operative store
   v. Super markets

C. Different sectors of Indian economy

- Agriculture and forestry
- Auto components
- Automobile
- Aviation
- Fintech and financial service
- Biotechnology
- Capital goods
- Chemical
- Construction
- Defense manufacturing
- Electronic systems
- Food Processing
- Healthcare
- IT (Information Technology)and BPM (Business Process management)
- Leather
- Media
- Medical Devices
- Metal Mining
- Oil and Gas
- Pharmaceuticals
- Ports and shipping
- Railways
- Renewable energy
- Retails and e-commerece
- Roads and highways
- Telecom
- Textiles and Apparels
- Thermal Power
- Tourism and hospitality

D. Supply chain models of companies

1. Amazon

The goal of the company is to create the best customer centric country. The chain of events that takes place after a customer places an order from Amazon are as follows

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 04, April-2021

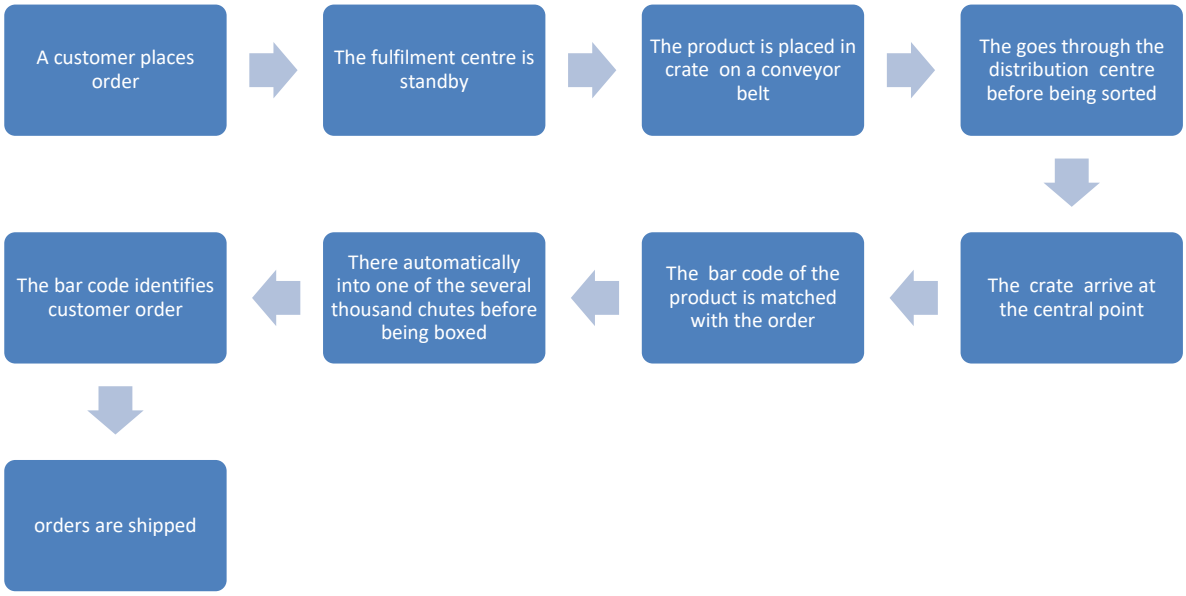


Fig 1:- SCM flow of amazon

There are six crucial elements that come into play for the logistics of supply chains in amazon

[i]. Warehousing
This is a critical part of amazon supply chain. It has around 175 fulfillment centers across the globe. These fulfillment centers are very sophisticated inventory to meet the demand across the globe. In amazon efficiency is very important so to maximize efficiency it divided into 5 unique storage area.

[ii]. Delivery
On time and fast delivery is one of the main focus of amazon.

Amazon offers a host of delivery options:

- Prime (1-2 hours, Same Day, One-Day and Two-Day Options)
- Hub & Locker Delivery
- Amazon Key (In-Home / In-Car / In-Garage Options)
- Amazon Scout Robot (Currently in test mode)
- Hub Counter with local retail partners
- Designated Day Delivery (to group orders together)
- No-Rush Shipping (6-day shipping with a discount incentive)
- Release Date Deliveries (Pre-order Option)

[iii]. Technology
One of the main reason for the success of amazon is the lavish use of technology. Amazon extensively uses robots in it fulfillment centres for various activities. Right amazon drone delivery system is being tested for 30 minutes delivery time. This wouls require a amazon landing mat and close distance to amazon delivery centres



Fig 2 :- process flow of amazon delivery services

Amazon seller
There is a service called Amazon seller in which you can create an account on Amazon and sell your products on Amazon. This process requires four steps.

Step1:- create an account an Amazon seller central account in sell.amazon.com. There are two types of account you can make in Amazon seller account. Professional and individual account. Professional account will cost $39.99 per month and $0.99per item for individual account. Individual account is much more basic account

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 04, April-2021

which lets you sell your own manufactured products but the professional account has benefits like fulfillment by Amazon, customize shipping rates, advertising, deals and coupon, automated pricing tools, Amazon sell coach, market place web services.

Step 2 :- The products must be listed on Amazon. you can either resell already existing products or sell your own products on Amazon. To sell a new product on Amazon you will need a product identifier that is a bar code – UPC, GTIN, ISBN or EAN.

STEP 3 :- The seller have to find a way to fulfill customer orders. They can use to fulfill their order. The customer using Amazon fulfillment by Amazon service will be given delivery, customer services, and distribution by Amazon. The customers selling on their own have to deliver return and distribute orders on their own.

Step4 :- Amazon will pay you on every 2 weeks of the month on eligible sales proceeds. The seller should become familiar with Amazon fee schedule.

2. Myntra

Myntra is an online shopping retailer selling fashion brands and casual life style products. It is the retailer which popularized coupon. It is currently ding 8000 transactions daily and over 4 lakh visits and the network has expanded to over 1,200 towns and cities across India endorsing over 350 national and international brands. It is operational in 12 major cities and has a paln of serving over 70 of it customer via their own logistics.

Some of the facilities myntra provides.

- It offers 24 hours delivery in most cities and free shipping within India. It can also deliver internationally to all major countries.
- Myntra invest heavily on IT sector to have a good customer experience. Plus it has a good digital marketing strategy.
- As myntra deals with merchandising procurement is important for them. Finished goods received from various brands are stored at warehouses and then they delivered to the customers.

Myntra uses third party couries to outbound their logistics. It is the company's responsibilities to the third party courier which will help it track its lead time and try and reduce it

The sourcing details of the company goes as follows

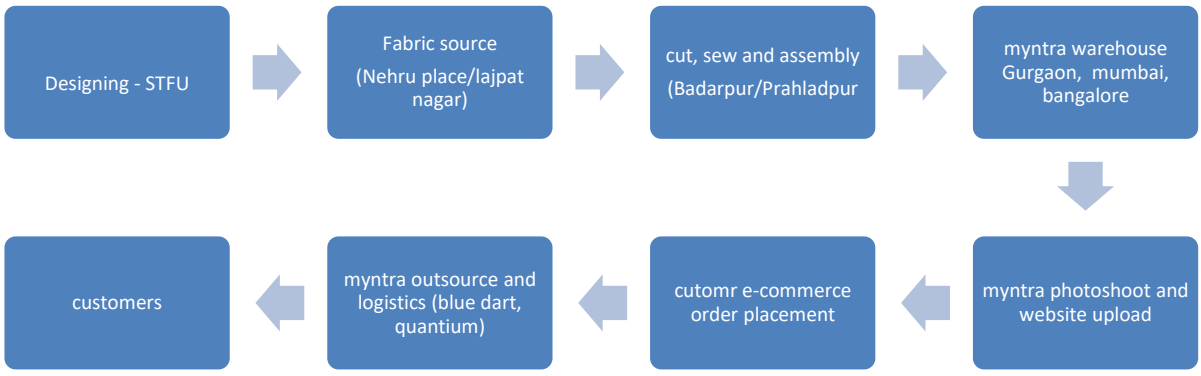


Fig 3:- Myntra supply chain flow chart

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 04, April-2021

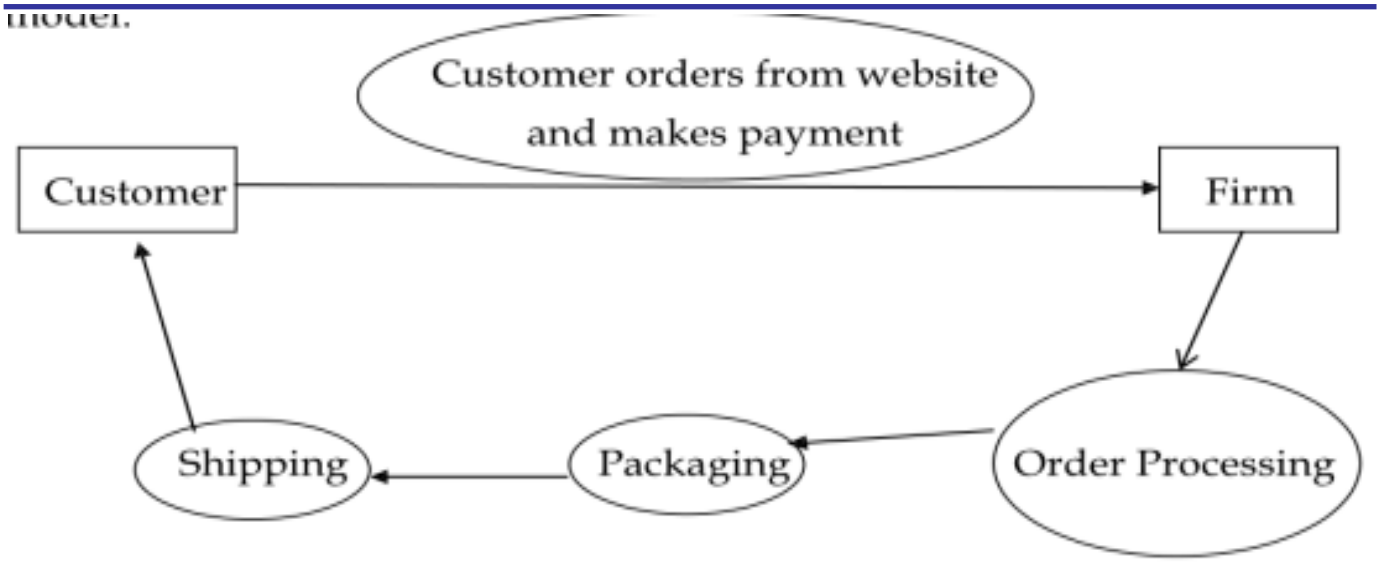


Fig 4:- Value chain analysis of myntra

3. Flipkart

Flipkart Advantage Shipment (FA)- available at flipkart fulfillment center

Non Flipkart Advantage- products ordered are available at marketplace sellers registered with flipkart.

ORDERS AT FA- After the order, the availability of the product is checked at the inventory.

Picklist with the detail of the product and its location is created and assigned to the associate. Product is packed as per the guidelines, the bill and label is pasted on the product

ORDERS AT NON-FA- the seller receives the order request and packs the product for shipment, first mile team leader creates a trip sheet and assigns a first mile field executive who goes to the seller location and verifies, picks and brings the shipment to the first mile team leader.

He bags the shipments as local, zonal, third party logistics, surface and national.

The shipments are transferred to the nearest mother hub. Mother hub receives the FA shipment from fulfillment centers and Non-FA shipments from first mile hubs, here the shipments are segregated into- Primary Segregation and Secondary Segregation.

Primary Segregation- the shipments are scanned and segregated on the basis of station basis .

Delivery Partner- Ekart

Station wise Dangerous Goods Same day and Next day Delivery(SDD AND NDD)

Secondary Segregation- the shipments are segregated on the basis of Destination Hub/COC code wise

1. NDD/SDD Normal shipments to respective DH
2. DG Shipments to respective MH
3. Third party logistic shipments according to vendors

Shipments are then scanned, weighed and sealed in white bags and stuck with bills and details these bags are then shipped to the line haul. It hands over the bag to the transport center team.

Shipments at Delivery hub- shipments are segregated and sorted in pin code wise or group wise. Arun sheet is generated and assigned to the FE, FE leaves the hub and visits customer premises for delivery, completes the payments transaction if any, updates the run sheet in the app and hands over the shipment

4. DHL (Dalsey, Hillblom & Lynn)

DHL is one of the world's leading delivery, supplier, packaging and Logistics Company of the world. DHL supplies compressive logistics solutions. They have four units of businesses they are :- DHL Express, Blue Dart express, DHL global forwarding, DHL supply chain. Even though DHL doesn't have an elaborate online site the reason we are studying the principles of supply chain of this company is because of its success and experience in establishing a global supply chain and logistics Excellency.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 04, April-2021

This will also make the retailer aware about the products and services retailers can use in order to take the help of DHL further for their businesses
Here we shall analyze various aspects of the company that make up the DHL supply chain.

[i]. The products and express that DHL provides.

- DHL document express:- The good which doesn't need to go through customs.
- DHL worldwide package express:- goods with commercial value that might go through customs inspection.
- Jumbo box and jumbo junior:- jumbo box is goods under 25 kg and jumbo junior is
  goods under 10kg. this services delivers goods anywhere in the world with one fee
- Import express:- This service allows you to import any goods around the world. It is a single window solution for delivery
- Fashion first:- this service is specialized for textile industry which need easy custom clearance and quicker delivery time
- Fast forward:- this is a service for transferring heavy goods 20kg or heavier. It has an online platform called trade automated service (TAS) which provides comprehensive trade and customs information of various country.
- DHL airlines:- it provides airlines services with different bases around the world.

[ii]. Intergraded logistics

There is a flow of ownership of the goods from finished product to customer which involves a lot of intermediate flow. This is made success with proper inventory flow and information flow.

- Inventory flow:- it involves distribution and procurement. The goods are weight and checked for condition and depending on the type of delivery it is packed

- Information flow:- there has to be a good information flow of the packages in its flow of inventory. There are two major types of information flow DHL operates :-
  Coordinate flow:- they try to match the location of the customer with the products and services to be performed and the estimated time it will require.
  Operation flow:- this involves directing various operation from a centralized flow of information. Operational flow deals with order management, order processing, distribution operation, inventory management, transportation and shipping, procurement.

- Forecasting:- based on years of experience historical data about the time required and the operations performed forecasting is done.

[iii]. Supply chain

DHL manages an international supply chain. The characteristics of its supply chain is such that it give clear picture to an upstream supply chain for faster decision making, the supply chain is more flexible which reduces lead time and inventory cost.

[iv]. Logistics services offered by DHL

- Reverse logistics:- this service helps customers plan , implement and control the flow of materials and information for the best possible outcome for their businesses. It involves installation of finished products in customer site, managing products on return, exchange of products.

- Service logistics:- this service involves replacement part delivery to and from customer according to the conditions given in the warranty.

- Inbound manufacturing:- this is a complete end to end logistics service for the purpose of management of inventory and labor to facilitate flow of material.

- Distribution to store management:- this service helps in creating adaptable supply chain.

Extended supply chain services

DHL provides other supply chain services

- Order management
- Call centre management
- Global inventory management
- Consolidated billing service
- Freight and customs services

They also provide quality management and product management services.

DHL solutions:- they provide logistics and supply chain solutions for a whole network. It involves all the process from counseling and planning to administration and execution. This is for different retailers.

DHL supply chain corporate information solution:- this provides warehouse and transportation services along with whole supply chain for clients.

[v]. Warehouse infrastructure network

DHL provides completely integrated warehousing facilities with wide supply chain.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 04, April-2021

- There are multi-user centers which provide a network of multi-user warehouse
- There are strategic part centers (SPC) Offering 1,2 and 4 hour fulfillment time.

- There are express logistics centre (ELC) which offers order processing, out source repair facility, custom find assembly, knitting services.

[vi]. W arehousing solutions of DHL

Warehouse management system (WMS) :- this manages all the records, actions, receipt and all other things that are required for the functioning of the warehouse.

Strategic inventory management (SIM):- this is a solution create specially for urgent or faster orders within 2 or 4 hours.

Direct express inventory management (DEI):-
This is a solution for customer so they can centralize their stocks in one warehouse and use express delivery components the next day.

Repair inventory management (RPI):- they provide a solution for the return, repair of the broken parts and moving them in and out of warehouse
Inventory optimization:- it basically focuses on the following things for the management of its industries :-

- Supplier management
- Expenditing
- Order replenishment
- Demand forecasting
- Safety stock setting
- Order pipeline monitoring
- Execess stock management

They use specialized softwares to carry out the above function

5. Assam Bio Refinery Private Limited (ABRPL)

ABRPL is in a joint venture with Numaligargh Refinary Limited, Chempolis OY and Fortum from Finland. This company aims to reduces the dependency on fossil fuel. So in order to produce petroleum there is a need of ethanol which is also a part of crude oil. As we know one the biggest expenditure of the country is the importation of crude oil. Therefore this company aims to produce bio-ethanol, acetic acid and furfural alcohol from bamboo biomass. Bamboo is a product which naturally grows in states like Assam, Nagaland, Arunachal Pradesh and Meghalaya. So here is an example of a supply chain that ABRPL is trying to establish with the help of local growers which will uplift the community of local farmers with help of a good app based supply chain. This will save millions of rupees and could potentially lower fuel price.

The stakeholders in this business are:
a) Bamboo grower
b) Harvesting and Interim Transporters
c) Local Level Entrepreneurs (LLE) for chipping bamboo
d) Transporter

[i]. Vendor registration system
Since the whole process is app based and to maintain the supply chain it is important to register all the bamboo growers, harvesters, LLE and transporter in the system.

There are different ways different vendors can register

a. Bamboo grower
   - They can visit the nearest ABRPL authorized store with required documents
   - Through an app or website
   - Through kioks put up by ABRPL

b. LLE
   - The responsibility of registration of LLE lies with ABRPL supply chain
   - They have to visit the nearest ABRPL authorized store with required documents.
   - The responsibility of registration of transporter lies with ABRPL supply chain
   - They have to visit the nearest ABRPL authorized store with required documents.

c. Harvesting group cum interim transporters
   - The responsibility of registration of transporter lies with ABRPL supply chain
   - They have to visit the nearest ABRPL authorized store with required documents.

All these vendors will have to fill a form filling in required details and produce the required documents. Then a 10 digit alpha numeric code will be generated which will be specific the vendor. They all will follow a specific protocol while verification.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 04, April-2021



Fig 5:- registration process

[ii]. Materials requirements planning

- ABRPL will have 14 days of chipped bamboo inventory in-house.
- Since ABRPL supply chain is Supply Constraint, material requirement planning will be done based on the plant production There has to be good amount of information sharing between ABRPL and LLE. 3 months tentative demand must be shared at the fourth week of each month. The requirements are calculated based on minimum dispatch quantity of chip, bamboo cost, transportation cost and inventory visibility.

plan, in-house inventory and in-transit inventory

- Below the flow chart of material planning is given

There also has to information sharing between ABRPL and bamboo growers . the communication can take place with the help of the app. They need to confirm the supply mode.

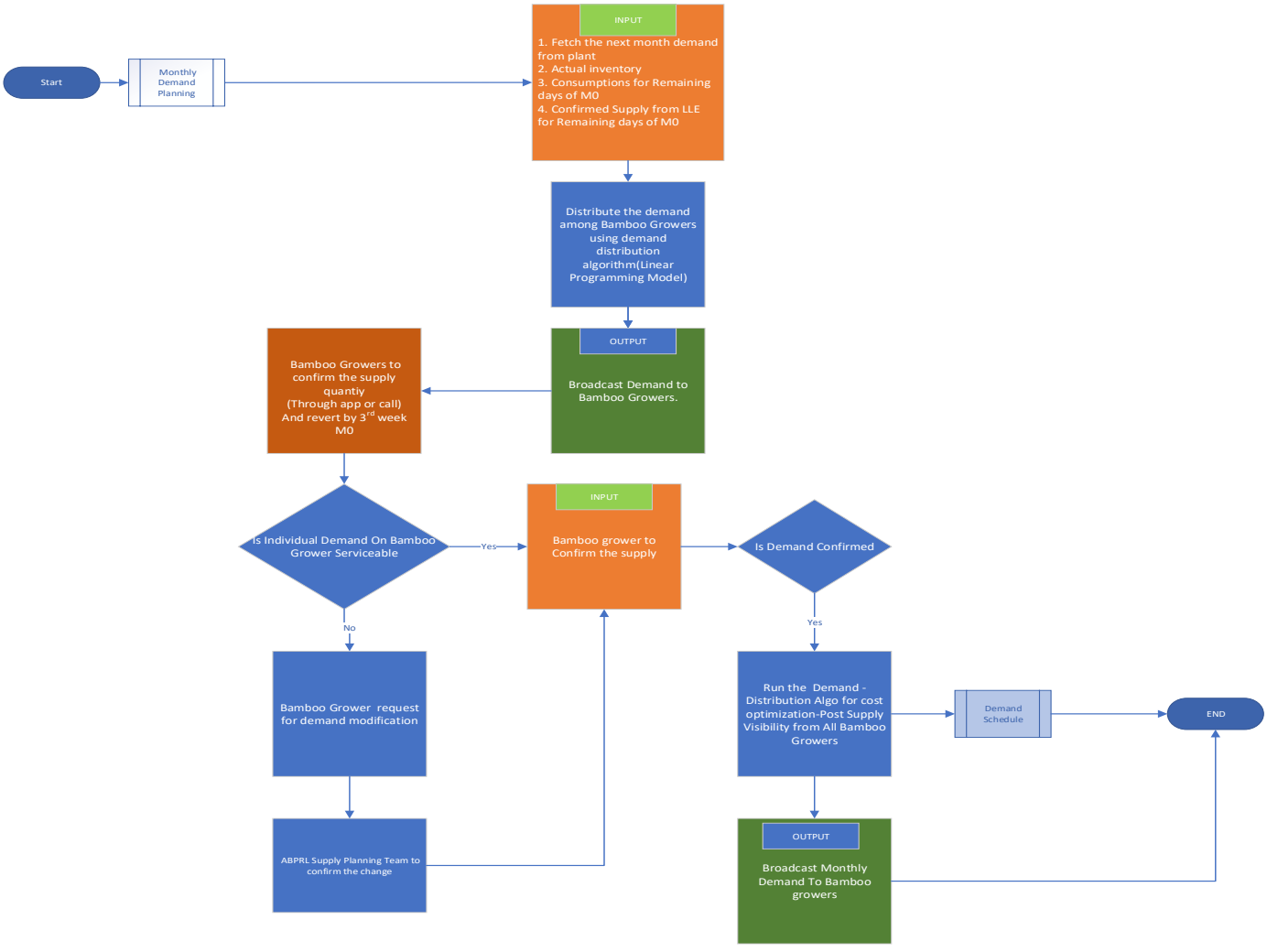


Fig 6:- Material management flow chart

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 04, April-2021

[iii]. scheduling

a. Between ABRPL and LLE
Here weekly delivery plan has to be shared with LLE with delivery date which will depend on the frequency of material supply. LLE can confirm, reject propose within 2 days. On reaching ABRPL they will confirm the moisture content.

b. Between LLE and farmer
Material delivery can take place by collection through various harvesting group designed by LLE or a simple delivery to ABRPL. The delivery quantity, transporter will be confirmed by ABRPL, they would be informed by a call and text. Material will be received with sipping id and bamboo grower id. On reaching LLE they shall confirm the moisture and weight and GR will be generated.

c. Bamboo growers and harvesting group
The harvesting group will confirm and pickup the harvest as and when directed by ABRPL. They shall feed all the information to the system. It will be confirmed by the bamboo growers ABRPL with information related to the weight and moisture content.

d. Direct supply to bamboo grower to ABRPL
It is possible for the bamboo growers to directly supply to ABRPL in the case of adversities. They have to request in the portal and on survey of the weight and moisture the bamboos will be accepted. GR receipt will be produced in the ABRPL gate.

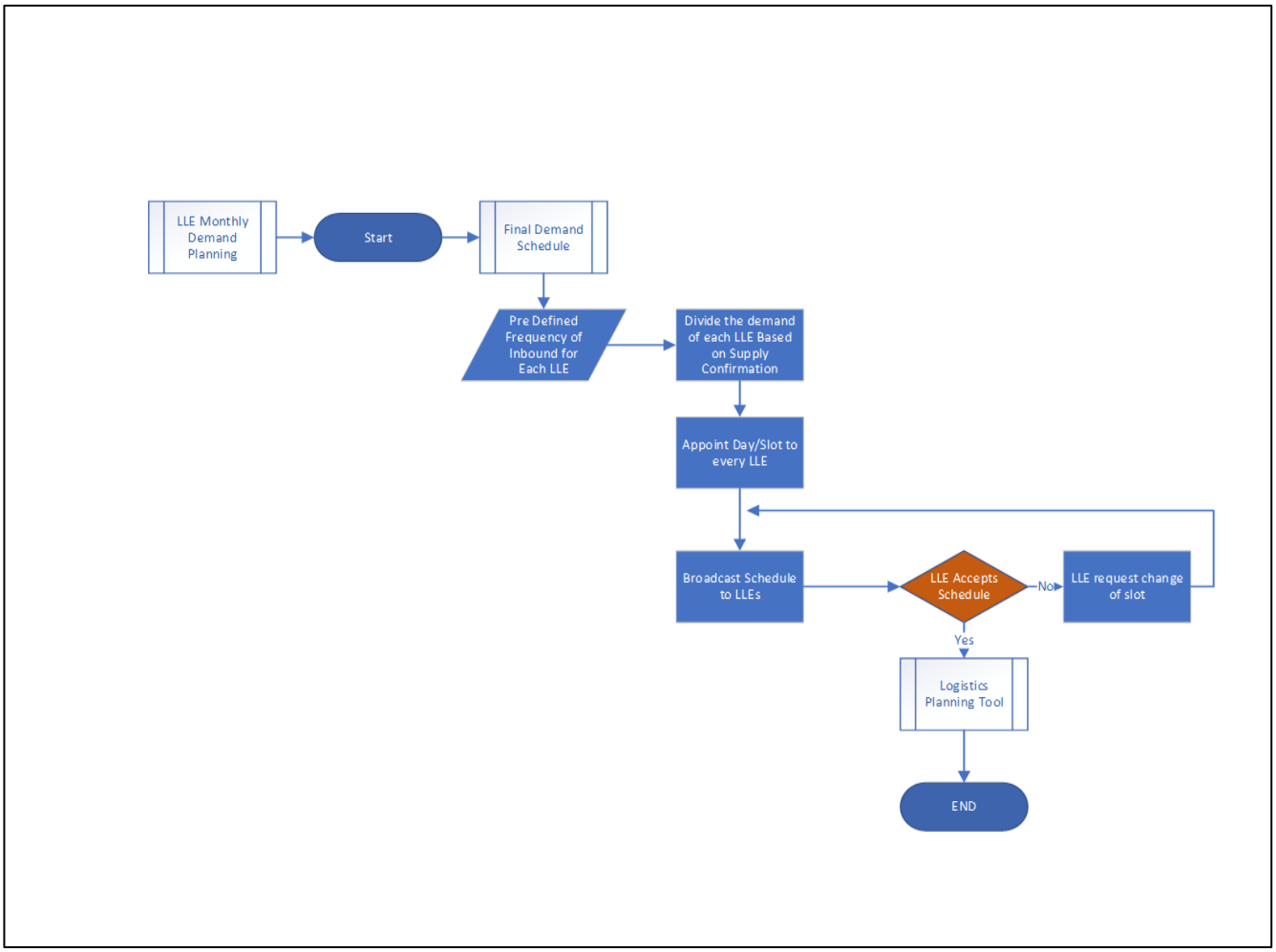


Fig 6:- Scheduling flow chart

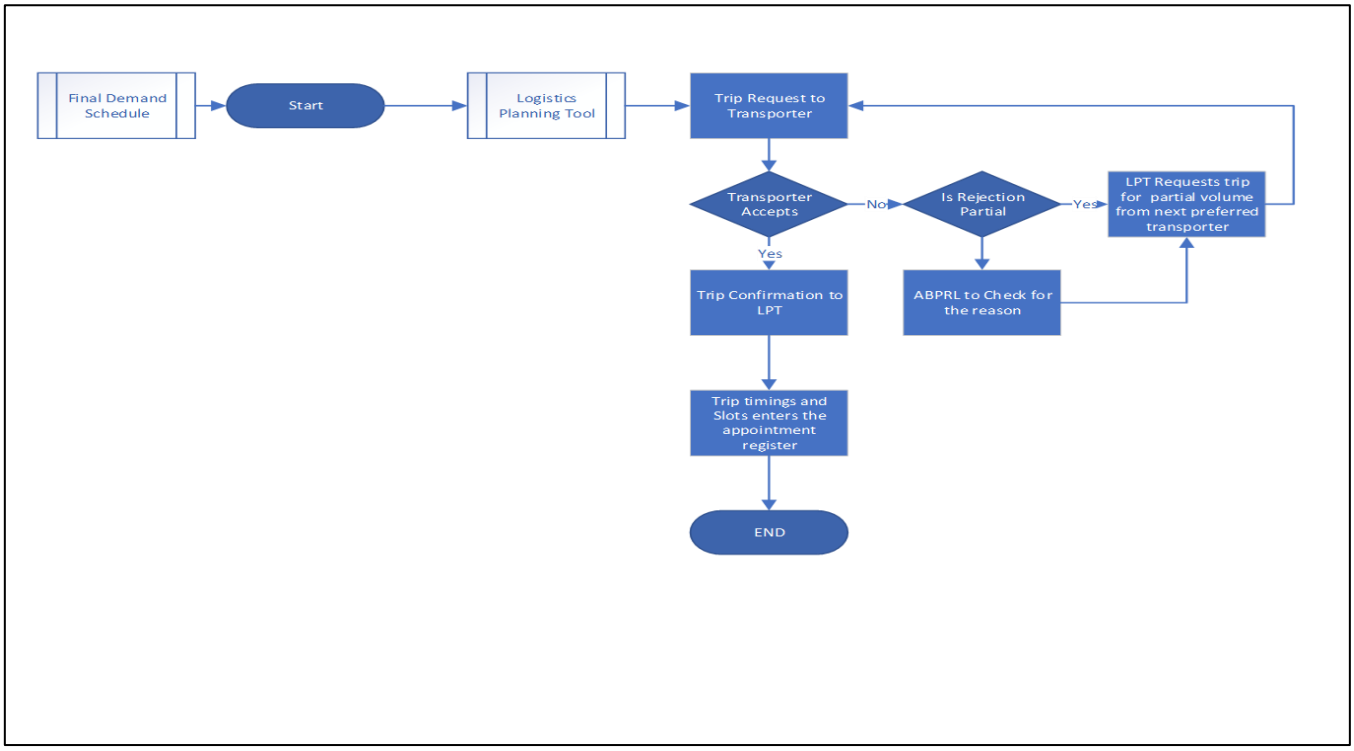
(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 04, April-2021

[iv]. Logistics management

a. Business distribution strategy will be done by ABRPL and the availability of the fleet, selection of transporter based on the agreed cost

A database has to be maintained of the transporter route wise and with settled rates. ABRPL will be deciding the business distribution for the individual routes. The selection of the transporter will be based on the rates and business distribution. Based on the weekly plan of the LLE, notification will be sent to the respective transporter. If the transporter accepts the request the shipping id will be shared to the LLE. If it is rejected it will be sent t the next preferred transporter.



b. Uber like model between routes where the fleet availability and demand is high

An uber like app has to be developed showing the available transporter and once accepted it can be tracked using GPS and the available transporter is seen on the near vicinity. The trip will go to the transporter with the cheapest rates. If the transporter rejects it will go to the next nearest transporter.

c. Online bidding for selected routes

There has to be a platform where the routes can be made available for bidding by ABRPL.

Transporter can get route information and submit the bid

[v]. material management at LLE

The harvest group will deliver the raw bamboo then they are weighted along with their shipping id and vendor code. Then they will be allotted a batch number pertaining to that day. Moisture content will be recorded and then FIFO will be maintained while chipping. Then the data will be stored by ABRPL against delivery id and respective LLE

[vi]. Payment system

All payments has to be direct bank to bank transfer into designated account of ABRPL.

- Bamboo grower

  Payment will be provided in credit cycle by ABRPL. The weight and the moisture content will be recorded by video analytics. Payment will paid as royalty component. The cost and rates will be decided by ABRPL. The bamboo growers will visit the bamboo growers as per schedule calculate the weight and moisture content and update it in the app.

  Weightment can also be done by LLE. Here all the weight should be taken by weightbridge and the data should be entered in the system.

- LLE

  Payment has to be provided within 7 days. System should have penalty and incentives. ABRPL should have the flexibility to change the payment cycles. The cost will depend on the proceesing cost agreed by ABRPL.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 04, April-2021

- Harvesting cum interim harvesting group
  The payment will be processed in 7 days. The cost will be based on the harvesting cost and transportation cost.

- Transporter
  The payment should be made within 7 days. The payment will be based on the video analytics made by the weighbridge. The rates will be fixed before hand.

III. RESULT

Our main objective of our research is to help the other retailers if they are willing to start e-retailing or if they wanted to get their products online and how to take help or establish a supply chain strategy which is going to aid or help in their business. In this paper we first search through the various kinds of retailers and markets existing in India. Then went through companies like myntra, Amazon, flipkart, DHL and ABRPL to take an example of how the supply chain works out for different companies. As a result we shall analyze the different supply chain strategies and try to bring out the best possible solutions for different businesses.

- Become a seller in Amazon. Amazon provides all the required facilities to sell products online, which is going to take care most of the supply chain needs if you a small retailer and creating and maintaining supply chain is not manageable by the retailer.
- If a retailer wants to outsource their supply chain they can refer to DHL supply chain which has excellent services in their supply chain which retailers can make use of to aid their business.
- The secret to the success of Amazon is the very sophisticated fulfillment centers. They have state of the art facilities in these fulfillment centers. They sort and distribute more a million packages per year. The use of robotic shelves, distribution tray. Their whole supply chain is customer focused which is one of the main reasons why it succeeded.
- The customer centric approach adopted by Amazon is one of their main reason why they succeeded. The question one should ask is why the customers would prefer shopping online. The answer will be cheaper price and less of hassle to go and shop outside. It has also been surveyed that customer sometimes doesn't prefer shopping online is slower delivery time. To which Amazon made two different supply chain for 24 hours delivery and 2 hour delivery time.
- Faster delivery time can be achieved by strategically placing the warehouses and the fulfillment centers.
- Once listed as an Amazon seller small retail shops will be able to deliver item according to Amazon prime now delivery supply chain.
- The use of technology has aided the smooth running of supply chain. For example the use of robots in fulfillment centers of Amazon
- The most important thing to start a business online is to have a dynamic fulfillment centre. In the case of flipkart all the items from individual sellers move into fulfillment centers.
- DHL is one the oldest and most fully equipped logistics company in the world. A lot of companies outsource their logistics and supply chain needs to DHL. It also provides supply chain solutions for smaller businesses. In thus paper we have seen the products that are available in DHL from logistics to transportations to inventory management and all other supply chain needs.
- ABRPL is excellent example of how a government company created a supply chain to fulfill it need and also uplift the farmers with a product which is produced from a locally grown plant-bamboo.
- One of the characteristics that you see common in all the supply chains and even DHL stretches heavily on it proper flow of information. This ensures the management of information in the supply chain.
- For transportation the companies with largest capital keep their own vehicles. Other that owning your fleet it is also possible to outsource it. Companies like DHL provide those services. We can also follow the model designed by ABRPL where we can register the vehicles to the company and through an app based we can request for a drop if the vehicle rejects the proposal it will go to the next nearest transporter in tis way we can reduce the cost of outsourcing and buy your own fleet.

IV. CONCLUSION

As our final year b-tech project we tried to research the market and their problems with supply chain and the current scenario status. We have survey the supply chain strategies and of four different companies who is trying to establish e commerce in various ways and government company which is trying to uplift the community along with creating solutions for the market. We have also studied the products of DHL which is an world renowned supply chain and logistics company. Many companies have the products and services of DHL as their third party logistic company. All these research was done with one aim in our mind and that to provide solutions for business which are planning on venturing into the e-retail business to keep up with the ever changing times.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 04, April-2021

V. ACKNOWLEDGEMENT

First of all, we would like to express my special gratitude to our university, Delhi Technological University, the Vice Chancellor, Professor Yogesh Singh sir and to our professor who guided us through the project, Professor N. Yuvraj sir, for providing us the necessary help. We also thank our friends and colleagues for their contribution and advice. We owe the completion of this paper within the limited time to all of them. The research the we did for the project has increased or knowledge to a great extent and has made us more informed on the topic.

The project has made us more knowledgeable, and the value of the knowledge we received from the project is beyond the grades.

REFERENCES

[1]. Kumar, Pravin. Industrial Engineering And Management. Pearson, (n.d.)

[2]. Nagpal, A. (2017) *'Challenges Faced By Indian Organized Retail Outlets: A Description Study Of Delhi/NCR'*, International Journal Of Management And Applied Science [Internet], Available from: http://www.iraj.in/journal/journal_file/journal_pdf/14-407-151142216922-27.pdf

[3]. https://www.economicsdiscussion.net/retailer/types-of-retailers/31832. (n.d.) 09 January 2021.

[4]. Balakrishnan, M. (2018) *'Status And Scope Of E-commerce In Agribusiness In India'*, International Research Journal Of Managment And Commerce [Internet], 14. Available from: https://www.researchgate.net/publication/329528325_Status_and_scope_of_e-commerce_in_agribusiness_in_India [09 January 2021].

[5]. Panigrahi, A., and Joshi, V. (2016) *'Application Of E-commerce In Indian Textile Industry'*, International Journal On Textile Engineering And Processing [Internet], Available from: https://www.researchgate.net/publication/323414054_Application_of_E-Commerce_in_Indian_Textile_Industry [12 January 2021]

[6]. *https://www.investindia.gov.in/sectors*. (n.d.) 16 January 2021.

[7]. *https://www.business2community.com/brandviews/xsellco/how-the-amazon-supply-chain-works-02359957#:~:text=The%20Amazon%20supply%20chain%20simply,ensure%20a%20smoot.* (n.d.) 19 January 2021.

[8]. (n.d.) *https://www.youtube.com/watch?v=itsGY_1ryCA. (Accessed 18 January 2021)*

[9]. (n.d.) *https://tinuiti.com/blog/amazon/amazon-supply-chain/*. (Accessed 22 January 2021)

[10]. (n.d.) *https://www.youtube.com/watch?v=3_MHpuixtOY*. (Accessed 23 January 2021)

[11]. (n.d.) *https://www.youtube.com/watch?v=Okq0xvlagXo&t=1776s*. (Accessed 23 January 2021)

[12]. Ahsan, Sayed Shamikh. "*Challenges In Logistics And Supply Chain Management At DHL Express*." (n.d.) https://www.researchgate.net/publication/342001425_U_Challenges_in_Logistics_and_Supply_Chain_Management_At_DHL_Express. Web 27 January 2021.

[13]. *http://hl.lookbookhq.com/ao_product_ecommerce/research-report_ecommerce-supply-chain-overcoming-growing-pains?utm_source=Eloqua&utm_medium=E-Mail&utm_campaign=WelcomeSeriesEmail19&utm_term=WelcomeSeries1&utm_content=Nurturing&sfid=7012S000001F7hDQASv_ref=WelcomeSeriesEmail1&elq_mid=5151&elq_cid=661698* (n.d.) 30 January 2021

[14]. (n.d.) Available at: http://dhl.lookbookhq.com/ao_product_transportation/research-report_logistics-transport-evolution-the-road-ahead?utm_source=Eloqua&utm_medium=E-Mail&utm_campaign=WelcomeSeriesEmail19&utm_term=WelcomeSeries1&utm_content=Nurturing&sfid=7012S000001F7hDQASnu_ref%3DWelcomeSeriesEmail1&elq_mid=5151&elq_cid=661698 (Accessed 01 February 2021)

[15]. (n.d.) Available at: http://dhl.lookbookhq.com/ao_product_ecommerce/infographic_ecommerce-supply-chain-three-key-insights?utm_source=Eloqua&utm_medium=E-Mail&utm_campaign=WelcomeSeriesEmail19&utm_term=WelcomeSeries1&utm_content=Nurturing&sfid=7012S000001F7hDQASv_ref=WelcomeSeriesEmail1&elq_mid=5151&elq_cid=661698 (Accessed 05 February 2021)

[16]. (n.d.) Available at: http://dhl.lookbookhq.com/ao_product_ecommerce/whitepaper_logistics-trend-radar?utm_source=Eloqua&utm_medium=E-Mail&utm_campaign=WelcomeSeriesEmail19&utm_term=WelcomeSeries1&utm_content=Nurturing&sfid=7012S000001F7hDQASv_ref=WelcomeSeriesEmail1&elq_mid=5151&elq_cid=661698 (Accessed 07 February 2021)

[17]. (n.d.) Available at: http://dhl.lookbookhq.com/ao_product_ecommerce/article_four-trends-shaping-future-ecommerce?utm_source=Eloqua&utm_medium=E-Mail&utm_campaign=WelcomeSeriesEmail19&utm_term=WelcomeSeries1&utm_content=Nurturing&sfid=7012S000001F7hDQASv_ref=WelcomeSeriesEmail1&elq_mid=5151&elq_cid=661698 (Accessed 07 February 2021)

[18]. (n.d.) Available at: http://dhl.lookbookhq.com/ao_product_ecommerce/article_behind-the-hidden-complexity-of-ecommerce-fulfillment?utm_source=Eloqua&utm_medium=E-Mail&utm_campaign=WelcomeSeriesEmail19&utm_term=WelcomeSeries1&utm_content=Nurturing&sfid=7012S000001F7hDQASv_ref=WelcomeSeriesEmail1&elq_mid=5151&elq_cid=661698 (Accessed 07 February 2021)

[19]. (n.d.) Available at: http://dhl.lookbookhq.com/ao_product_transportation/video_managed-transport?utm_source=Eloqua&utm_medium=E-Mail&utm_campaign=WelcomeSeriesEmail19&utm_term=WelcomeSeries1&utm_content=Nurturing&sfid=7012S000001F7hDQASnu_ref%3DWelcomeSeriesEmail1&elq_mid=5151&elq_cid=661698 (Accessed 07 February 2021)

[20]. (n.d.) Available at: http://dhl.lookbookhq.com/ao_product_transportation/infographic_logistics-transport-evolution-road-ahead?utm_source=Eloqua&utm_medium=E-Mail&utm_campaign=WelcomeSeriesEmail19&utm_term=WelcomeSeries1&utm_content=Nurturing&sfid=7012S000001F7hDQASnu_ref%3DWelcomeSeriesEmail1&elq_mid=5151&elq_cid=661698 (Accessed 07 February 2021)

[21]. (n.d.) Available at: http://dhl.lookbookhq.com/ao_product_transportation/article_four-ways-to-increase-value-of-tms?utm_source=Eloqua&utm_medium=E-Mail&utm_campaign=WelcomeSeriesEmail19&utm_term=WelcomeSeries1&utm_content=Nurturing&sfid=7012S000001F7hDQASnu_ref%3DWelcomeSeriesEmail1&elq_mid=5151&elq_cid=661698 (Accessed 09 February 2021)

[22]. (n.d.) Available at: http://dhl.lookbookhq.com/ao_product_transportation/article_three-transport-innovations-that-deliver-roi?utm_source=Eloqua&utm_medium=E-Mail&utm_campaign=WelcomeSeriesEmail19&utm_term=WelcomeSeries1&utm_content=Nurturing&sfid=7012S000001F7hDQASnu_ref%3DWelcomeSeriesEmail1&elq_mid=5151&elq_cid=661698 (Accessed 09 February 2021)

[23]. (n.d.) Available at: http://dhl.lookbookhq.com/ao_product_transportation/webinar_reimagining-the-future-of-transport?utm_source=Eloqua&utm_medium=E-Mail&utm_campaign=WelcomeSeriesEmail19&utm_term=WelcomeSeries1&utm_content=Nurturing&sfid=7012S000001F7hDQASnu_ref%3DWelcomeSeriesEmail1&elq_mid=5151&elq_cid=661698 (Accessed 10 February 2021)

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 04, April-2021

[24]. (n.d.) Available at: http://dhl.lookbookhq.com/ao_product_transportation/webinar_reimagining-the-future-of-transport?utm_source=Eloqua&utm_medium=E-Mail&utm_campaign=WelcomeSeriesEmail19&utm_term=WelcomeSeries1&utm_content=Nurturing&sfid=7012S000001F7hDQASnu_ref%3DWelcomeSeriesEmail1&elq_mid=5151&elq_cid=661698 (Accessed 11 February 2021)

[25]. (n.d.) Available at: https://www.youtube.com/watch?v=xN5scjazLE8 (Accessed 12 February 2021)

[26]. (n.d.) Available at: https://abrpl.azurewebsites.net/ (Accessed 13 February 2021)

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 6, Issue 2, ISSN No. 2455-2143, Pages 62-66
Published Online June 2021 in IJEAST (http://www.ijeast.com)



SUSTAINABLE MANUFACTURING - THE PAST, PRESENT AND FUTURE

Aryaman Sharma
U.G. Student, Computer Engineering
Delhi Technological University, New Delhi, India

Abstract: - Sustainable manufacturing has become of utmost importance during these times of strained supply chain logistics and availability as highlighted by the Covid-19 pandemic. Organizations, both privately and publicly owned have realized the positive relation of economic progress and sustainable manufacturing which have made them imbibe and implement such practices. Newer technologies, efficient processes and waste reducing methods etc. have become key in any manufacturing process. This trend will continue increasing with time for production to remain economical and regular. Almost all research on this topic has focused on a particular timeline of sustainable manufacturing. The purpose of this paper is to explore the wider timeline of events that sustainable manufacturing has gone through in order to reach this importance.

Keywords - Sustainable manufacturing, Covid-19, Economical, Economic progress., Efficient.

I. INTRODUCTION

The term "Sustainable Development" was first coined in the Brundtland Commission Report in 1987 titled - Our Common Future: Report of the World Commission on Environment and Development. According to the United Nations, sustainable development can be defined as "Development that meets the needs of the present without compromising the ability of future generations to meet their own needs." [1]

Companies and main players in the Supply-Demand chain are often the key entities who have to propagate and practice sustainable development goals in order for a significant change to come [2]. The Sustainable Development goals are often imbibed by the companies in their Corporate Social Responsibilities [3]. The majority of manufacturing facilities are only spectators in the global energy supply chain. They keep track of total energy use and prices connected with available sources, but they frequently ignore the controllable aspects within operations that enable consumption management that adds value to the company. This is where the concept of Sustainable Manufacturing comes into play. The United States Environmental Agency defines Sustainable Manufacturing as "the creation of manufactured products through economically-sound processes that minimize negative impacts while conserving energy and natural resources." Sustainable Manufacturing is a very important niche in the study of Sustainable Development. Especially during the COVID 19 pandemic, the role of manufacturing in providing critical supplies to various sectors such as Healthcare, Food, Energy etc. has been amplified.

For example, the Center for Disease Control (CDC) which is the national public health agency of the United States and CISA (Cybersecurity and Infrastructure Security Agency) have identified critical staff belonging to various critical sectors and published various guides and safety manuals in order to safely and effectively continue manufacturing processes with minimum exposure of the people to the virus [4]. Recognizing the importance of sustainable manufacturing processes and care for the environment, the United Nations Conference on Trade and Development which was founded in 1964 initiated a Sustainable Manufacturing and Environmental Pollution (SMEP) program that had started in 2019 and will end in 2024. Along with the UNCTAD, Foreign Commonwealth and Development Office (FCDO), Department of Environment and Geography at the University of York. Due to weak environmental regulations, poor manufacturing technologies and insufficient funding, the small industries worldwide leading to increased environmental and ecological damage. Therefore, setting up SMEP would help micro-, small- and medium sized Enterprises (MSME's) to gain technological know-how in order to prevent the environmental degradation taking place in Sub-Saharan African and South Asian countries. [5]. The SMEP scoping study will use internationally recognized exposure-response relationships (ERRs) and thresholds/guidelines for human health and environmental degradation to provide a semi-quantitative assessment of damage from key industrial activities in order to translate levels of pollution into exposure and damage. [6]

The COVID-19 situation provides an opportunity to develop a versatile and resilient manufacturing framework that will ensure the production process's economic and social sustainability in the event of any such global pandemics and emergencies in the future. Therefore, it is imperative that Manufacturers, Governments and people themselves understand the importance of this situation and act on it

International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 6, Issue 2, ISSN No. 2455-2143, Pages 62-66
Published Online June 2021 in IJEAST (http://www.ijeast.com)



collectively and comprehensively as no single party can drive the need in for Sustainable Manufacturing by themselves.

The aim of writing this research paper is to explore the gradual shift that has taken place in the Manufacturing Sector over the years and the emerging trends that are expected to have an impact on the manufacturing practices in the foreseeable future. We will see how from practicing Unsustainable manufacturing processes to imbibing Sustainable ones, integrating lean and green, environmentally friendly techniques, the manufacturing sector has changed.

Previously, writers and researchers has discussed a specific timeline in the timeline of Sustainable Manufacturing. However, very few to no research has focused to see the complete timeline and compare how it is growing in prevalence and demand due to various ecological changes that have been going on for the past few decades such as Climate Change, shift to more sustainable methods of development etc. Therefore, there is a need to fill this void and explore the evolution of Sustainable Manufacturing through time.

Another reason for this issue to be researched further is due to the recent public outcry for environmental conservation and protection that has increased manifolds over the last few decades. In a 2014 study conducted by the Clinton Global Initiative and Microsoft, it was found that about 76% of millennials surveyed were more focused on the environment [7]. Therefore, it makes a strong business case for the manufacturers to be sensitive to the concerns of this important customer segment who would like to see the products to be environmentally concerned. With the advent of technology, Reduction, reuse, recycling and remanufacturing are connected to the environmental improvements in the production processes. The manufacturing system considers zero emits (i.e., closed loops) to be an industrial ecosystem, requiring the reuse of waste or by-products in the manufacturing system. Thus, the production of zero emissions requires pollution prevention (e.g., replacement) and waste reuse capabilities.

Flexible production also requires material flexibility, and production equipment that can accommodate changes in material flows. This in turn can help to improve sustainability while at the same time maintain competitiveness. For example, more effective, recyclable packaging can improve the sustainability of packaging.

II. HISTORY

Historically, the idea of sustainable manufacturing and consumption can be traced to an essay written by Thomas Malthus (1766-1834) in 1798. He was an English cleric and scholar. His essay, “An Essay on the Principle of Population”, he argued that while the population of humanity would keep growing exponentially over the coming decades, the food production levels would grow arithmetically [8]. He had taken the case of the United Kingdom for this study. Although this idea was presented about two hundred years ago, this idea is very relevant even today. The philosophy of sustainability has been synonymous with dystopian rhetoric since Thomas Malthus identified rapid population growth as a harbinger of humanity's doom.

Then around 1997, John Elkington defined the term “Triple Bottom Line”. This approach requires that companies must measure their social and environmental aspect in addition to their financial performance other than their profit-making activities. Such early attempts at defining and conceptualization Sustainable Manufacturing helped modern day studies and thinkers to plan for the future and create new techniques to prevent overutilization of resources [9]

Historically, it is seen that the rate of consumption of resources always increases multifold with increasing population thanks to new developments and better medical services. People nowadays have consumed resources alarmingly fast, which is definitely not sustainable, because the Earth we live in cannot regenerate the materials at that rate. In comparison, world metals manufacturing grew steeply between 1950 and 2005, oil consumption was eight times higher and natural gas consumption was 14 times higher [10]. Seeing the increased consumption of minerals and energy resources, lean manufacturing processes with minimal harm to the environment must be taken in order to ensure that the resources are efficiently and effectively used. Market leaders and investors have started to accept sustainability in recent years because of its many upside potentials.

III. THE PRESENT STATE AND INCREASING RELEVANCE

If we see the global energy consumption and CO_2 emissions in the year 2005, it is clear that the manufacturing sector contributed most in both the above-mentioned areas, contributing about 33% and 38 % respectively [11]. Many multinational organizations today are attracted towards sustainable manufacturing processes mainly due to reduced cost and lesser resource wastages [12]. Various multinational companies spanning various niches of the economy like Technology, Homeware, Cosmetic etc. have profited and made billions by shifting to more sustainable, lean and greener manufacturing processes. The general public, as stated in the introduction, also support such initiatives taken by them and therefore do not mind paying extra in order to propagate the message of sustainability [13]. One vital tendency found is the economic motivation for lowering fees and enhancing productivity. Many cases are also influenced by way of the environment and society (government incentives, sector/ customer pressure, environmental rules

International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 6, Issue 2, ISSN No. 2455-2143, Pages 62-66
Published Online June 2021 in IJEAST (http://www.ijeast.com)

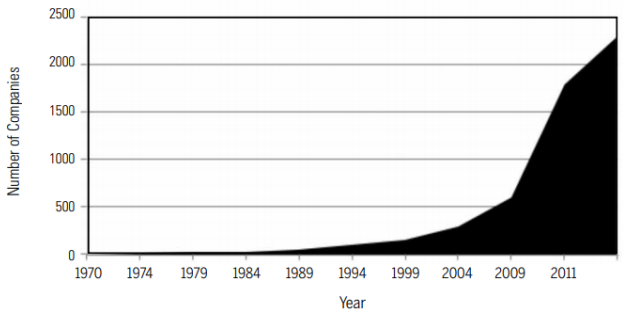
and new standards) ensuing in businesses integrating extra and more sustainability in their company method and ethics. This recollects again the three interrelated dimensions of sustainability (people, planet and profit), the 3 P's, and shows that companies are primarily transferring towards sustainability through compliance to societal adjustments and requirements [14].

For example, in the Aerospace industry, Sustainable Manufacturing has taken on a more important and significant role. Use of cutting-edge materials critically affects both real-life and natural problems. The meaning of materials science and development in Airplanes couldn't possibly be more significant. The materials used as a piece of aviation and aerospace designs and car applications are essential to its successful plan, structure and working. Materials have their effect all through the whole life cycle of airplane, beginning from the plan period of their item advancement to their finish of-life removal

Various steps can be taken in order to reduce the wastage of materials, energy and wasteful byproducts formed by production of various components of an aircraft [15].

The below provided graph gives the growth in Sustainability to various companies over the past few decades. It is clear that it has picked up pace during recent years.

Figure 1. Rise in Importance of Sustainability to Companies



(SOURCE: Success Paths to Sustainable Manufacturing, 2014, https://oneill.indiana.edu/doc/research/sustainability-2014.pdf)

IV. FUTURE OUTLOOK AND COURSE OF ACTION

Manufacturing's past success has been built on a willingness to take chances, and the future of industrial development will be no different. Manufacturing's past success has been built on a willingness to take chances, and the future of industrial development will be no different. The Fourth Industrial will be driven by those who invest in radical innovation and promote the adoption of low-carbon, resource efficient technology at scale, innovative sustainable businesses and collaborative collaborations models. Manufacturers who develop ultra-efficient processes and assist in the development of resilient supply chains will be in a prime position to protect themselv es against resource shortages and escalating commodity cost s while also collecting latent economic potential [2].

Innovative technology of today and tomorrow will make sustainable manufacturing simpler than ever before. Companies can now manufacture parts on-site rather than have them shipped from afar thanks to advances in 3D printing (or additive manufacturing). Systems and products can now interact with one another thanks to nanotechnology and the Internet of Things (IoT). This, when combined with machine learning and artificial intelligence, allows computers to automatically classify and implement the most efficient manufacturing methods. In the past, these minor but significant changes would have necessitated hundreds of hours of data analysis and calculations. They can now be finished in the blink of an eye. [16].

V. CHALLENGES

The traditional role of product design in manufacturing firms has been to ensure that the product achieves its desired goals based on consumer tastes, taking performance, reliability, ergonomics, and aesthetics into account. With a greater emphasis on sustainable production, the product designer now has the added burden of taking into account the environmental effects of his or her decisions. The materials and methods used have a direct impact on the final product. As a result, the designer must first familiarize himself or herself with the various issues surrounding sustainability. The environmental effect is heavily influenced by the materials and methods used [17].

Another challenge to embracing Sustainable Manufacturing is the pricing aspect of it. Due to competition, the price of the products produced must be low and at the same time must be able to produce a profit for the product to be economically viable for the company. Many industries do not want to lose out on their customers and do not ant to procure

If substantial results are to be achieved, some aspects of improving efficiency, such as changing factory layout, require time and physical space. Interviewees at the improvement sites proposed that a temporary "buffer" of time be created by using evenings or weekends, adding additional resources to the area, or temporarily increasing the number of people working in the area. Time by working evenings and weekends, committing additional resources to the field, or temporarily growing stock to 'make room' for improvement activities without affecting production requisites [18].

However, there is still a long way to go in order to properly implement these practices on a much wider scale. Knowledge with regards to sustainable production field is fragmented

however unified theories, typically accepted frameworks and fashions are developing.

VI. CONCLUSIONS

This paper explores the timeline of Sustainable Manufacturing processes, its implementation through time, how companies can and have implemented it and the challenges in fully implementing such techniques. Climate change and sustainability will fundamentally change manufacturing as each company revisits where plants are located, how materials are sourced and shipped and how tradeoffs in goods are produces. Productivity, efficiency, tracking and visibility will be more important than ever. Various challenges such as efficiency, pricing and need for large scale investing in newer technologies etc. hamper the progress of sustainable manufacturing.

VI. REFERENCES

[1] Darroch, J. (2017, September 13). What is Sustainable Development and Why Does it Matter? HuffPost. https://www.huffpost.com/entry/what-is-sustainable-devel_b_11966082

[2] The Fourth Industrial Revolution: what it means and how to respond. (2016, January 14). World Economic Forum. https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/

[3] Companies and sustainable development | Insee. (n.d.). Insee.Fr. Retrieved June 15, 2021, from https://www.insee.fr/en/metadonnees/source/serie/s1227

[4]https://www.cisa.gov/sites/default/files/publications/Version_3.0_CISA_Guidance_on_Essential_Critical_Infrastructure_Workers_1.pdf

[5] Sustainable Manufacturing and Environmental Pollution (SMEP) Programme. (2021). Sustainable Manufacturing and Environmental Pollution (SMEP) Programme. https://www.smepprogramme.org/s/

[6] united nations sustainable. (n.d.). Unctad.Org. https://unctad.org/project/sustainable-manufacturing-and-environmental-pollution-smep

[7] millennials survey. (n.d.). Msnbc.Com. https://www.msnbc.com/morning-joe/millennials-environment-climate-change-msna291876

[8] Malthus, T. (1998). An Essay on the Principle of Population. Www.Esp.Org. http://www.esp.org/books/malthus/population/malthus.pdf

[9] Miller, K. (2020, December 8). The Triple Bottom Line: What It Is & Why It's Important. Business Insights - Blog. https://online.hbs.edu/blog/post/what-is-the-triple-bottom-line

[10] Mineral Supply and Demand into the 21st Century. (2007, January). pubs.usgs.Gov. Retrieved May 10, 2021, from https://pubs.usgs.gov/circ/2007/1294/reports/paper9.pdf

[11] World on the Edge - Energy Data - Earth Policy Institute. Earth-policy.Org. Retrieved May 10, 2021, from http://www.earth-policy.org/datacenter/pdf/book_wote_energy_oil.pdf

[12] Sustainable manufacturing: Trends and Research Challenges. (2012). Journal of Trends in the Development of Machinery and Associated Technology. Published. https://www.tmt.unze.ba/zbornik/TMT2012Journal/01.pdf

[13] Rahimifard, S., Seow, Y., & Childs, T. (2010). Minimizing Embodied Product Energy to support energy efficient manufacturing. CIRP Annals, 59(1), 25–28. doi:10.1016/j.cirp.2010.03.048

[14] billion dollar sustainability. (2016). Theguardian.Com. https://www.theguardian.com/sustainable-business/2016/jan/02/billion-dollar-companies-sustainability-green-giants-tesla-chipotle-ikea-nike-toyota-whole-foods

[15] M. Despeisse , F. Mbaye , P.D. Ball & A. Levers (2012) The emergence of sustainable manufacturing practices, Production Planning & Control: The Management of Operations, 23:5, 354-376, DOI: 10.1080/09537287.2011.555425

[16] Applying lean in Aerospace manufacturing. (2017). Researchgate.Net. https://www.researchgate.net/publication/320476138_Applying_lean_in_aerospace_manufacturing

[17] How 3D printing is revolutionizing manufacturing. (2019, January 4). Hpe.Com. https://www.hpe.com/us/en/insights/articles/how-



3d-printing-is-revolutionizing-manufacturing-1901.html

[18] @inproceedings{Nambiar2010ChallengesIS, title={Challenges in Sustainable Manufacturing},Author={A. N. Nambiar},year={2010}}

[19] Implementing Lean in aerospace challenging the assumptions and understanding the challenges. (2003, December 1). ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/S0166497203000816

Inorganica Chimica Acta 527 (2021) 120580

Contents lists available at ScienceDirect

Inorganica Chimica Acta

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

Synthesis, structural characterization and binding ability of A_2B cobalt(III) corroles with pyridine

Jyoti [a], Natalia Fridman [b], Anil Kumar [a,*], Sudhir G. Warkar [a]

[a] Department of Applied Chemistry, Delhi Technological University, Delhi 110042, India
[b] Schulich Faculty of Chemistry, Technion-Israel Institute of Technology, Haifa, Israel

ARTICLE INFO

Keywords:
Cobalt
Corroles
Binding constant
NMR

ABSTRACT

A series of five A_2B type bis-pyridine ligated cobalt complexes $Co[(p\text{-}NO_2Ph)_2RCor](py)_2$, has been synthesized where R belongs to different electron withdrawing or electron donating substituent at *meso* position for different complexes. The synthesized complexes were fully characterized through ^1H NMR, HRMS, UV–Visible and single crystal X-ray Spectroscopy. The effect of *meso* substituents were studied through NMR spectroscopy. The single crystal X-ray structure was obtained for the p-trifluoromethylphenyl substituents (**Co-2**) located at C_{10}-*meso* position and indicated that cobalt corroles exist in Co(III) form. In these compounds, the axial pyridines were observed to be labile, and the complexes were almost entirely dissociated in the five coordinated cobalt corrole monopyridine complexes in dilute solutions such as in dichloromethane. Mono- and bis- pyridine ligation behavior was studied by UV–Visible spectroscopy. These complexes were also examined in different non-aqueous media such as dichloromethane, benzonitrile, and pyridine. The effect of substituent behavior present at *meso* position was also tested through the Hill equation. This provided the fact that the lower the electron density on corrole macrocycle, more the tendency of the complex to bind the ligands axially as the highest binding constant was found for the **Co-1** complex, with more electronegative groups on *meso* position.

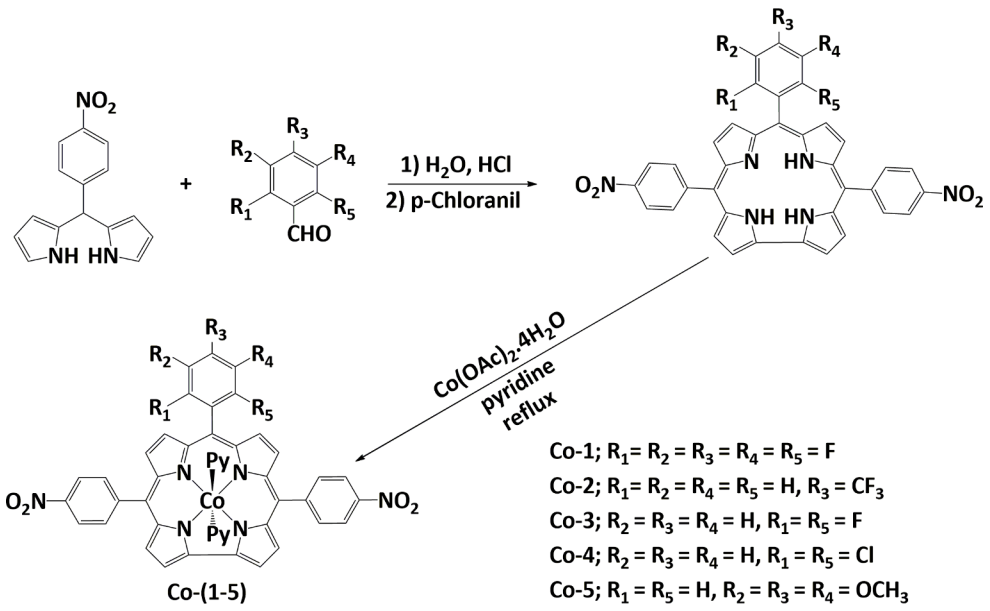
1. Introduction

Cobalt corrole complexes have been the subject of attention hence are the most often studied type of metallocorroles. This is because, in general cobalt complexes exist in Co(III) oxidation state but when co-ordinated in a corrole macrocycle, they exhibit a complex redox chemistry [1–5]. These complexes offers both, metal- and ring-centred reactions depending on the reaction parameters, such as type and number of substituents on *meso*-, β- and axial position [6–8]. This essential property has been assessed by the use of various techniques, e. g., UV–Visible, NMR, DFT, cyclic voltammetry and controlled-potential electrolysis. Due to all these features cobalt corrole complexes have been proved as a potential catalyst in a variety of reactions such as ORR [4,6–12], HER [13–17], OER [18], CO_2 reduction [19–21] and in various sensing applications where these are used to prepare different sensing materials [22–28].

The axially co-ordinating ligands which exert a huge effect on the electrochemical and redox properties of these metallocorroles have lately been focused on [7,20,29–34]. The cobalt corroles with axially ligating pyridine, DMSO, NH_3 or CO have evidenced this impact as these co-ordinating ligands cause not only changes in electron density at the metal ion, but also, in certain instances, a change in the oxidation state of both metal and ligand [29–31]. The ligand co-ordination with the central cobalt incorporated into the macrocyclic ligand is largely found solvent and concentration dependent. The anion interaction of a series of different cobalt A_3-triarylcorroles was recently reported, where among 11 different anions, namely CN^-, PF_6^-, BF_4^-, HSO_4^-, ClO_4^-, Br^-, I^-, Cl^-, OAc^-, F^-, and OTs^-, the only exception was CN^- that interacted with the neutral form of Co-corroles [35]. This led to the formation of mono- and bis-cyanocobalt complexes in DCM solution, the first example of air stable Co-Corroles with anionic axial ligands. Later, their chemical properties were discussed in comparison to mono-DMSO precursors, bis-pyridine analogues and the transient mixed-ligand adducts comprising one CN^- and one pyridine axial ligand [36]. For CO_2 conversion to CO, Sinha et. al. carried out comparative investigation of cobalt chelated superstructured corroles with PPh_3 and bis-pyridine present axially was also conducted where these complexes performed efficiently towards CO_2 reduction [20]. To further insight into the spectroscopic properties of cobalt corroles, we have synthesized a new series of A_2B type, bis-pyridine ligated complexes where two *meso*

* Corresponding author.
E-mail address: anil_kumar@dce.ac.in (A. Kumar).

https://doi.org/10.1016/j.ica.2021.120580
Received 17 May 2021; Received in revised form 7 August 2021; Accepted 13 August 2021
Available online 19 August 2021
0020-1693/© 2021 Elsevier B.V. All rights reserved.



Scheme 1. Synthetic route of A_2B cobalt corroles **Co–(1 to 5).**

positions, C_5 and C_{15}, are p-NO_2 phenyl substituted whereas substituents present at C_{10} position are either electron withdrawing or releasing group (Scheme 1). In this study, we investigated the impact of electron withdrawing or electron donating substituents on the binding constant of these cobalt complexes with pyridine in detail [37]. NMR spectroscopy for these complexes was also studied which revealed a strong shift in pyridine protons as the electron density increased on the corrole macrocycle, indicating that different substituents had a distinct effect on the overall structure of the complex. The X-ray crystallographic parameters and the chemical shift of β-protons inferred that cobalt is coordinated to a corrole macrocycle and present in low spin Co(III) complex.

2. Experimental section

2.1. Instrumentation

Nuclear magnetic resonance spectroscopy: ^1H and ^{19}F NMR spectra were recorded on a Bruker Avance spectrometer operating at 400 MHz for ^1H and 377 MHz for ^{19}F NMR. Chemical shifts are reported in ppm relative to residual solvent peaks of chloroform ($\delta = 7.26$ ppm) or benzene ($\delta = 7.16$ ppm) in the ^1H NMR spectra. The obtained spectra were evaluated by using MestRenova software.

Mass Spectrometry: High-resolution mass spectra (HRMS) of complexes were recorded by Maxis Impact HD Mass Spectrometer and APCI (atmospheric pressure chemical ionization) was used as a direct probe in either positive or negative mode.

UV–visible spectroscopy: Spectral measurements were carried out with the UV-1800 Shimadzu Spectrophotometer. Cuvettes made up of Quartz with optical path lengths of 1 cm were used.

2.2. Binding constant Determination:

Binding constant for pyridine was experimentally determined by UV–visible spectroscopic method for all the complexes. In this method, the titration of cobalt complexes solution with pyridine was monitored using UV–visible spectroscopy. Later the data was fitted in Hill equation (Eq. (1)) and the graph between concentration of pyridine vs absorbance was plotted [38,39].

$$log\left[\frac{(A_i - A_0)}{(A_f - A_i)} = log\mathrm{K} + \mathrm{n}log[\mathrm{L}]\right] \quad (1)$$

where A_0 is the initial absorbance when concentration of ligand [L] is zero, A_i is absorbance at a specific concentration of L and A_f signifies final absorbance value when only a fully ligated species of corrole complex remains in solution. logK values were obtained from intercept of the graph plotted between $\log((A_i - A_0)/(A_f - A_i))$ vs log [L] and is the binding constant for 'n' number of ligands coordinated axially.

Chemicals: All chemicals and solvents were used as purchased without further purification else otherwise noted. Silica gel (230–440 mesh) and neutral alumina were used for column chromatography. Benzonitrile (PhCN) purchased from Sigma-Aldrich chemical Co. was used. Pyridine (99%) and p-chloranil (98%), were obtained from Avra chemicals and dichloromethane (CH_2Cl_2) from Merck and used without further purification.

Synthesis: The same procedure for the preparation of all the desired five cobalt corroles (**Co-1 to 5**) was followed. To prepare the free base corroles, the procedure based on Gryko *et al.* [40] work was utilized and then insertion of cobalt metal was performed using $Co(OAc)_2{\cdot}4H_2O$ in pyridine (acting as solvent as well as coordinating ligand) to obtain the corresponding metallocorroles derivatives (Scheme 1) [41]. The free base synthesis involved dipyrromethane (1 mmol) and aromatic aldehyde (0.5 mmol) dissolved in 150 mL of CH_3OH in a round-bottom flask. Then 2.5 mL of HCl (36%) in 50 mL water was added to it and stirred at room temperature for 2 h. The mixture was extracted with $CHCl_3$, washed twice with water, and dried over Na_2SO_4. The organic layer was then filtered and diluted to a volume of 250 mL using $CHCl_3$. The mixture was then stirred overnight at room temperature after adding p-chloranil (1.5 mmol) as an oxidising agent to it. This reaction mixture was then chromatographed on a silica gel column using the DCM/Hexane as eluent. The green coloured free base corrole was obtained which further underwent metal insertion reaction and resulted in the desired cobalt corrole **Co–(1–5)** formation. For the metalation of free base (**1**–**5**) reaction, in a 100 mL round-bottom flask the synthesised free base corrole (19 mg) and $Co(OAc)_2{\cdot}4H_2O$ (33.61 mg) was dissolved in pyridine (15 mL) and the resulting solution was refluxed with stirring in an oil bath at 100 °C for 20–30 min. The completion of the reaction was monitored through TLC. Afterwards, the reaction mixture was cooled to room temperature and pyridine was evaporated on a rotary evaporator. The residue was dissolved in a minimum amount of DCM and passed

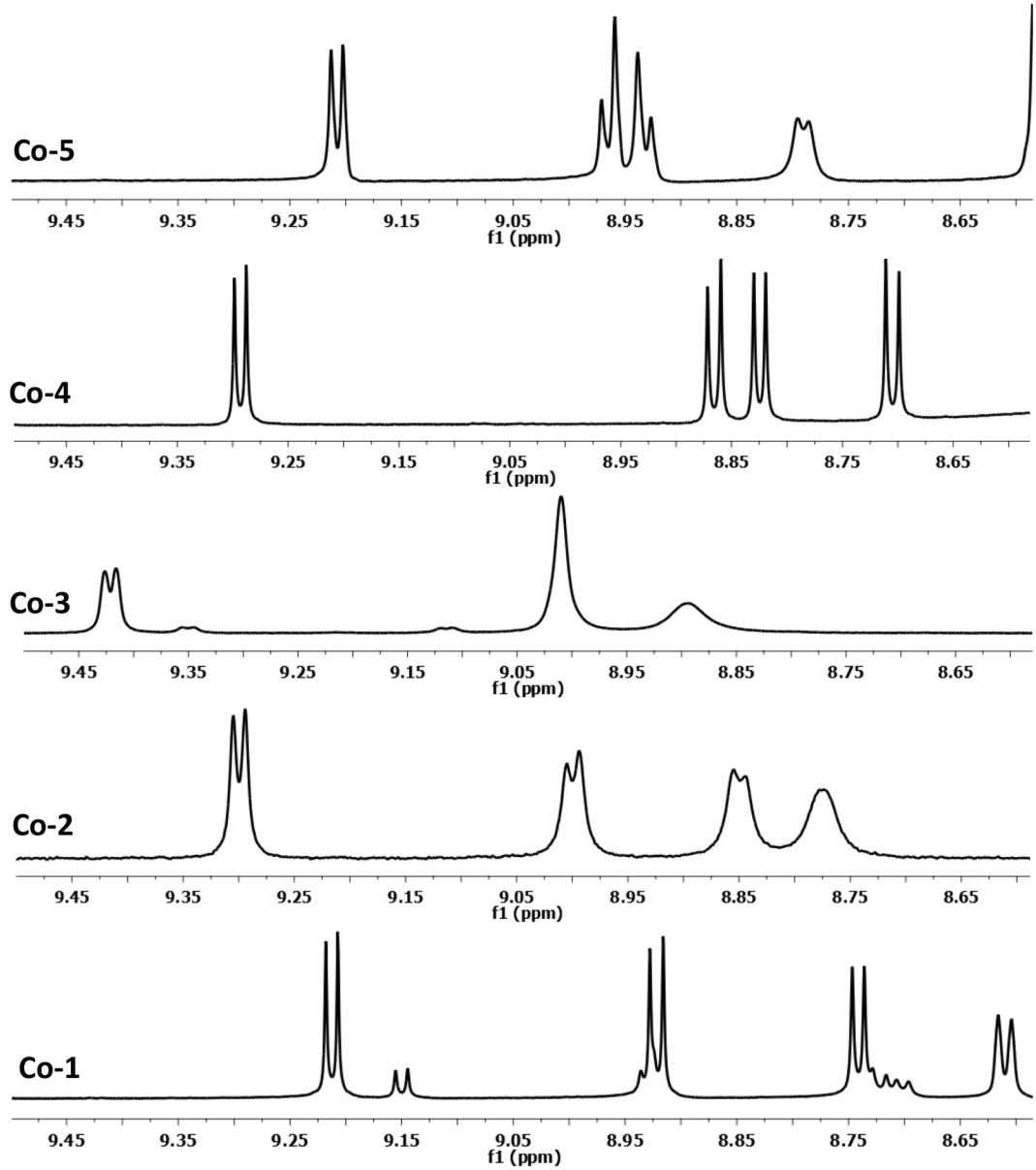


Fig. 1. ^{1}H NMR (room temperature, 400 MHz) spectra of **Co–(1 to 4) recorded in C_6D_6** and **Co-5** recorded in $CDCl_3$.

over the silica gel column. The mixture containing DCM, hexane as well as pyridine under different concentrations (described below for each corrole) was used as eluent.

Bis-pyridine [10-(2,3,4,5,6-*pentafluorophenyl*)-5,15-*bis*(4-*nitrophenyl*)*corrolato*] Cobalt(III) (Co-1) was obtained through silica gel chromatography with DCM:Hexane:Pyridine (2:8:0.1) as eluent in 79% yield (19.51 mg, 0.0198 mmol). UV–Visible [DCM + pyridine (0.1%); λ, nm]: 376 (50.4), 426 (49.3), 652 (0.305). ^{1}H NMR (400 MHz, C_6D_6) δ 9.21 (d, $^{3}J_{H,H}$ = 4.3 Hz, 2H, β-pyrrolic), 8.92 (d, $^{3}J_{H,H}$ = 4.8 Hz, 2H, β-pyrrolic), 8.74 (d, $^{3}J_{H,H}$ = 4.3 Hz, 2H, β-pyrrolic), 8.61 (d, $^{3}J_{H,H}$ = 4.7 Hz, 2H, β-pyrrolic), 8.04 (d, $^{3}J_{H,H}$ = 8.7 Hz, 4H, *o/m*-Ph), 7.85 (d, $^{3}J_{H,H}$ = 8.7 Hz, 4H, *o/m*-Ph), 4.58 (br s, 2H, *p*- H of pyridine), 3.95 (br s, 4H, *m*-H of pyridine), 1.33 (br s, 4H, *o*-H of pyridine overlapped with a solvent signal). ^{19}F NMR (377 MHz, C_6D_6) δ −139.37 (dd, $^{3}J_{F,F}$ = 26.2, 9.1 Hz, 2F, *o*-F of Ph), −153.77 (t, $^{3}J_{F,F}$ = 20.3 Hz, 1F, *p*-F of Ph), −162.86 (ddd, $^{3}J_{F,F}$ = 26.1, 22.3, 8.9 Hz, 2F, *m*-F of Ph). HRMS: Calcd for $C_{37}H_{16}CoF_5N_6O_4$ ([M−2py]$^{+\cdot}$) *m/z* 762.0485, *m/z* found 762.0504 ([M−2py]$^{+\cdot}$).

Bis-pyridine [10-(4-*trifluoromethylphenyl*)-5,15-*bis*(4-*nitrophenyl*)*corrolato*] Cobalt(III) (Co-2) was obtained through silica gel chromatography with DCM:Hexane:Pyridine (4:8:0.1) as eluent in 82% yield (20.42 mg, 0.0227 mmol). UV–Visible [DCM + pyridine (0.1%); λ, nm ($\varepsilon \times 10^{-3}$, L mol^{-1} cm^{-1})]: 370 (54.3), 421 (47.1), 562 (9.1), 644 (22.3). ^{1}H NMR (400 MHz, C_6D_6) δ 9.30 (d, $^{3}J_{H,H}$ = 3.8 Hz, 2H, β-pyrrolic), 9.00 (d, $^{3}J_{H,H}$ = 3.4 Hz, 2H, β-pyrrolic), 8.85 (d, $^{3}J_{H,H}$ = 2.7 Hz, 2H, β-pyrrolic), 8.78 (br s, 2H, β-pyrrolic), 8.24 (d, $^{3}J_{H,H}$ = 8.6 Hz, 4H, *o/m*-Ph), 8.18 (d, $^{3}J_{H,H}$ = 7.9 Hz, 2H, o/*m*-Ph), 8.10 (d, $^{3}J_{H,H}$ = 8.6 Hz, 4H, *o/m*-Ph), 7.74 (d, $^{3}J_{H,H}$ = 8.0 Hz, 2H, o/*m*-Ph), 4.82 (t, $^{3}J_{H,H}$ = 7.3 Hz, 2H, *p*-H of pyridine), 4.20 (br s, 4H, *m*-H of pyridine), 1.72 (br d, 4H, *o*-H of pyridine). ^{19}F NMR (377 MHz, C_6D_6) δ −61.64 (s, 3F, *p*-CF_3 of Ph). HRMS: Calcd for $C_{38}H_{20}CoF_3N_6O_4$ ([M−2py]$^{+\cdot}$) *m/z* 740.0830, *m/z* found 740.0817 ([M−2py]$^{+\cdot}$).

Bis-pyridine [10-(2,6-*difluorophenyl*)-5,15-*bis*(4-*nitrophenyl*)*corrolato*] Cobalt(III) (Co-3) was obtained through silica gel chromatography with DCM:Hexane:Pyridine (5:8:0.1) as eluent in 78% yield (19.63 mg, 0.0227 mmol). UV–Visible [DCM + pyridine (0.1%); λ, nm ($\varepsilon \times 10^{-3}$, L mol^{-1} cm^{-1})]]: 377 (88.1), 446 (68.3), 560 (19.0), 650 (54.1). ^{1}H NMR (400 MHz, C_6D_6) δ 9.42 (d, $^{3}J_{H,H}$ = 4.1 Hz, 2H,

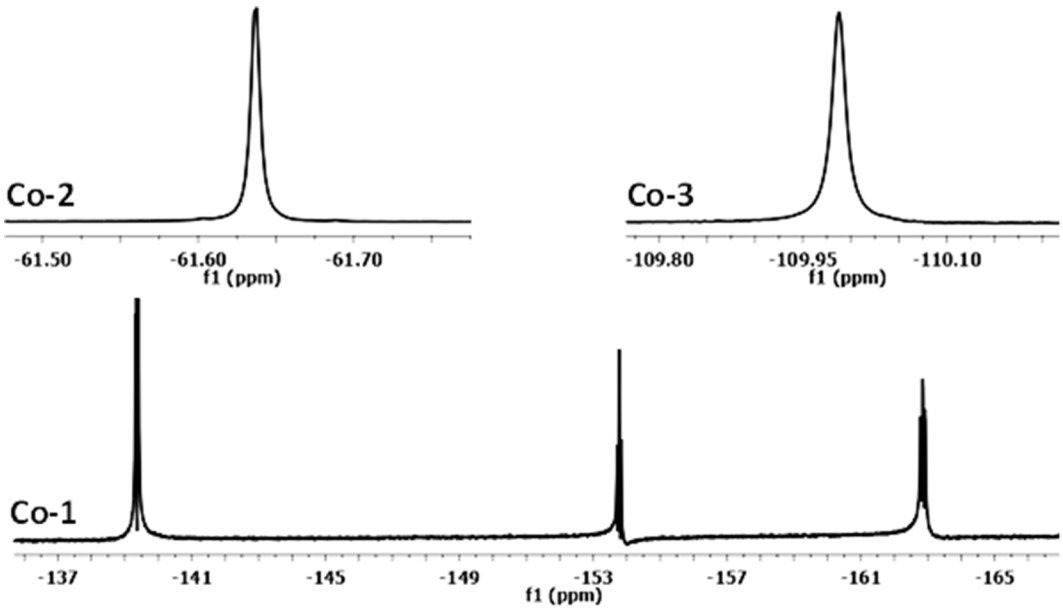


Fig. 2. ^{19}F NMR (room temperature, 377 MHz) spectra of **Co–(1 to 3)** recorded in C_6D_6.

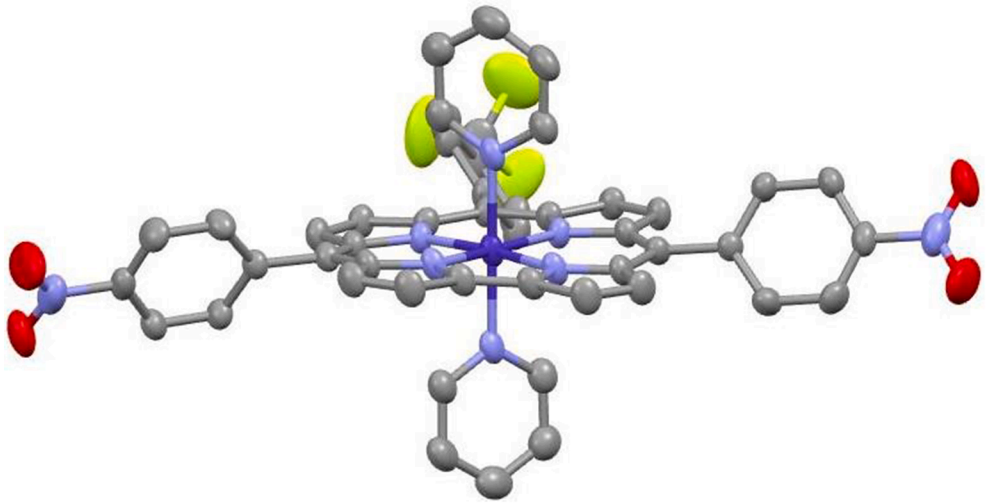


Fig. 3. X-ray crystal structure of complex **Co-2.** Thermal ellipsoids are shown at the 30% probability level.

β-pyrrolic), 9.01 (br s, 4H, *β*-pyrrolic), 8.90 (br s, 2H *m*-Ph), 8.22 (d, $^3J_{H,H}$ = 8.1 Hz, 4H, *o/m*-Ph), 8.04 (d, $^3J_{H,H}$ = 8.1 Hz, 4H, *o/m*-Ph), 7.09 – 7.02 (m, 3H *p*-Ph overlapped with a deuterated solvent residual signal), 4.83 (br s, 2H, *p*-H of pyridine), 4.24 (d, $^3J_{H,H}$ = 6.4 Hz, 4H, *m*-H of pyridine), 1.76 (br s, 4H, *o*-H of pyridine). ^{19}F NMR (377 MHz, C_6D_6) *δ* −109.99 (s, 2F, *o*-F of Ph). HRMS: Calcd for $C_{37}H_{19}CoF_2N_6O_4$ ($[M-2py]^{+\cdot}$) *m/z* 708.0768, *m/z* found 708.0754 ($[M-2py]^{+\cdot}$).

Bis-pyridine [10-(2,6-*dichlorophenyl*)-5,15-*bis*(4-*nitrophenyl*)*corrolato*] Cobalt(III) (Co-4) was obtained through silica gel chromatography with DCM:Hexane:Pyridine (5:8:0.1) as eluent in 81% yield (20.15 mg, 0.022 mmol). UV–Visible [DCM + pyridine (0.1%); *λ*, nm ($\varepsilon \times 10^{-3}$, L mol^{-1} cm^{-1})]: 378 (67.4), 419 (63.1), 649 (47.3). ^1H NMR (400 MHz, C_6D_6) *δ* 9.29 (d, $^3J_{H,H}$ = 4.3 Hz, 2H), 8.87 (d, $^3J_{H,H}$ = 4.8 Hz, 2H), 8.82 (d, $^3J_{H,H}$ = 4.3 Hz, 2H), 8.70 (d, $^3J_{H,H}$ = 4.8 Hz, 2H), 8.06 (d, $^3J_{H,H}$ = 8.7 Hz, 4H), 7.88 (d, $^3J_{H,H}$ = 8.7 Hz, 4H), 7.31 (d, $^3J_{H,H}$ = 8.2 Hz, 2H), 6.83 (t, 1H *p*-Ph, Ph overlapped with a deuterated solvent residual signal), 5.22 – 4.99 (m, 2H, *p*- H of pyridine), 4.08 (br s, 4H, *o*-H of pyridine), 2.01 (br, 4H, *m*-H of pyridine). HRMS: calculated for $C_{37}H_{19}Cl_2CoN_6O_4$ ($[M-2py]^{+\cdot}$) *m/z* 740.0177, *m/z* found 740.0208 ($[M-2py]^{+\cdot}$).

Bis-pyridine [10-(3,4,5-*trimethoxyphenyl*)-5,15-*bis*(4-*nitrophenyl*)*corrolato*] Cobalt(III) (Co-5) was obtained through silica gel chromatography with DCM:Hexane:Pyridine (3:8:0.1) as eluent in 79% yield (22.47 mg, 0.0244 mmol). UV–Visible [DCM + pyridine (0.1%); *λ*, nm ($\varepsilon \times 10^{-3}$, L mol^{-1} cm^{-1})]]: 376 (70.8), 426 (69.9), 570 (17.1), 652 (32.6). ^1H NMR (400 MHz, $CDCl_3$) *δ* 9.21 (d, $^3J_{H,H}$ = 4.3 Hz, 2H, *β*-pyrrolic), 8.95 (br m, 4H, *β*-pyrrolic), 8.79 (d, $^3J_{H,H}$ = 3.9 Hz, 2H, *β*-pyrrolic), 8.57 (d, $^3J_{H,H}$ = 8.6 Hz, 4H, *o/m*-Ph), 8.42 (d, $^3J_{H,H}$ = 8.6 Hz, 4H, *o/m*-Ph), 7.18 (s, 2H, *o*-Ph), 5.21 (s, 2H, *p*-H of pyridine), 4.06 (s, 3H, $-OCH_3$), 3.89 (s, 6H, $-OCH_3$), 1.56 (br s, overlapped with a H_2O solvent signal, 4H, *o*-H of pyridine). HRMS: Calcd for $C_{40}H_{27}CoN_6O_7$ ($[M-2py]^{+\cdot}$) *m/z* 762.1273, *m/z* found 762.1275 ($[M-2py]^{+\cdot}$).

2.3. Single crystal X-ray structure determination

Single crystals of $C_{48}H_{30}CoF_3N_8O_4$ (**Co-2**) were grown by slow evaporation of DCM solution of compound in heptane. A suitable crystal was selected and mounted on a diffractometer. The crystal was kept at

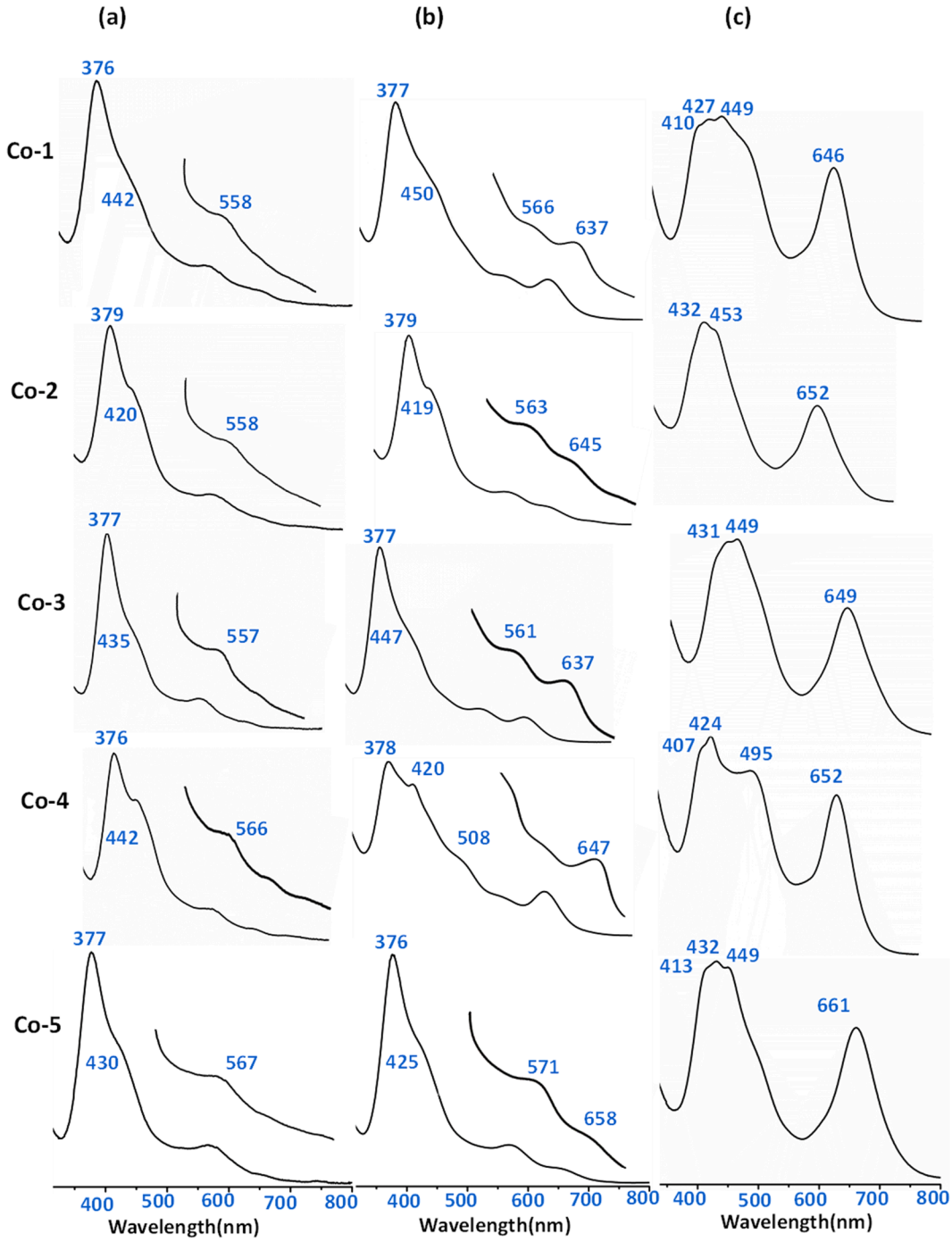


Fig. 4. Electronic absorption spectra of **Co–(1 to 5)** in (a) 10^{-6} M DCM (b) 10^{-4} M DCM (c) 10^{-5} pyridine solution.

200.15 K during data collection. Using Olex2 [42], the structure was solved with the olex2.solve [43] structure solution program using Charge Flipping and refined with the SHELXL [44] refinement package using Least Squares minimization. The structure consists of fluorine atoms: F1, F2, F3 disordered over two sites with occupancies 0.5:0.5; nitro groups: O1, O2 and O1A, O2A disordered over two sites with occupancies 0.9:0.1; carbon atoms: C12, C13, C15, C16 and C12A, C13A, C15A, C16A disordered over two sites with occupancies 0.91:0.9.

Crystal data. $C_{48}H_{30}CoF_3N_8O_4$ (M = 898.73 g/mol): monoclinic, space group $C2/c$ (no. 15), a = 18.5859(15) Å, b = 16.8267(13) Å, c = 13.3423(11) Å, β = 109.583(2)°, V = 3931.3(5) Å^3, Z = 4, T = 200.15 K, μ(MoKα) = 0.511 mm^{-1}, *Dcalc* = 1.518 g/cm^3, 15,364 reflections measured (3.356° ≤ 2Θ ≤ 50.278°), 3494 unique (R_{int} = 0.0768, R_{sigma} = 0.0627) which were used in all calculations. The final R_1 was 0.0512 (I > 2σ(I)) and wR_2 was 0.1488 (all data). **CCDC No. 2081897.**

3. Results and discussion

Synthesis. A series of cobalt corroles (**Co-1 to 5**) was synthesized in good yield following the procedure described in literature [40,41]. The reaction involved the corresponding free base corroles refluxed at 100 °C in the presence of Co(II) acetate in pyridine for 20–30 min (Scheme 1). The reaction was monitored through UV–Visible spectroscopy as well as with TLC and the appearance of strong Q-band confirmed the association of cobalt metal in the corrole macrocycle affording the six-coordinated bispyridine-cobalt-corroles. Once the reaction was

Table 1
UV–Vis spectral bands of **Co–(1–5)** in (a) 10^{-6} M DCM (b) 10^{-5} pyridine (c) 10^{-5} benzonitrile solution.

Solvent	Compound	λ nm ($\varepsilon \times 10^{-4}$ L mol^{-1} cm^{-1})			
		Soret region bands		Visible region bands	
Dichloromethane	**Co-1**	376 (21.9)	442 (11.3)		558 (4.2)
	Co-2	379 (26.0)	420 (18.3)		558 (3.4)
	Co-3	377 (20.9)	435 (9.9)		557 (3.8)
	Co-4	376 (23.4)	442 (17.6)		566 (4.1)
	Co-5	377 (19.1)	430 (10.6)		567 (5.5)
Pyridine	**Co-1**	410 (5.0)	427 (5.2)	449 (5.3)	646 (4.0)
	Co-2	432 (5.4)	453 (5.2)		652 (3.4)
	Co-3	431 (5.2)	449 (5.3)		649 (2.9)
	Co-4	407 (4.8)	424 (5.12)	495 (4.3)	652 (3.7)
	Co-5	413 (5.12)	432 (5.3)	449 (5.1)	661 (3.7)
Benzonitrile	**Co-1**	426 (6.4)	444 (6.2)		646 (4.1)
	Co-2	430 (6.1)	453 (5.5)		651 (2.7)
	Co-3	427 (7.0)	443 (6.8)		646 (4.5)
	Co-4	407 (6.1)	424 (6.4)	502 (5.0)	653 (4.1)
	Co-5	415 (6.6)	448 (5.7)	598 (1.8)	661 (3.2)

completed, the reaction mixture was cooled to room temperature, evaporated the solvent and the pure compound was extracted using column chromatography. For elution, ~1–5% pyridine was added with DCM/Hexane during the chromatographic separation. In the absence of it, impure compound with a variety of side products were produced. This phenomenon was noticed and shown to lead to C—C bond dimers. TLC, ^{1}H NMR (Figs. S1–S5) spectra and HRMS (Figs. S9–S14) were used to validate the purity of the green compounds produced. The X-ray crystals of bispyridine[10-(4-*trifluorophenyl*)-5,15-*bis*(4-*nitrophenyl*)*corrolato*] Cobalt(III) **(Co-2)** were also obtained in good quality which further confirmed the proof of composition and insertion of cobalt metal as well ligated pyridine.

NMR Spectroscopy. The ^{1}H NMR spectra for all the five compounds displayed sharp signals for β-pyrrolic protons in the 8.0–9.5 ppm range (Fig. 1). The signal due to pyridine protons, on the other hand, showed 4:4:2 of integrated peaks and were found strongly downfield shifted. For example, the pyridine protons of **Co-1** resonate at 4.58, 3.95 and 1.33 ppm for *p*-, *m*- and *o*-Hs respectively (Fig. S1). Similarly, the chemical shifts for **Co-2** were found at 4.64, 4.04 and 1.57 for *p*-, *m*- and *o*-Hs of pyridine respectively (Fig. S2). We do observe the variation in ^{1}H NMR of **Co-3** and **Co-4** complexes; 4 sharp doublets in the case of **Co-4** while one sharp doublet and 2 broad signals resonating at 9.01 and 8.90 ppm, for four proton and two proton of β-carbon of pyrrole, respectively. The diamagnetic ring current effect of corrole macrocycle has been proposed as the responsible factor for this extent of lower δ values [45–47]. ^{19}F NMR spectra of **Co–(1 to 3)** was also examined in C_6D_6 (Fig. 2). The *o*-, *m*- and *p*-F in Co-1 showed signals at −139.37, −162.86, −153.77 ppm respectively integrated in 2:2:1 (*-p*: *-m*: *-o* respectively), while **Co-2** and **Co-3** exhibited singlet peaks, respectively at −61.64 and −109.99 ppm. The well resolved spectra obtained for all the three complexes reflects that the pyridine ligands are coordinated symmetrically at axial positions, consistent with the octahedral geometry of complexes. The spectral range and nature of spectra of both the ^{1}H NMR and ^{19}F NMR is indicative of the existence of the diamagnetic Co(III) oxidation state of complexes [31,46]. Another conclusion is that the complexes are 6-coordinate with two pyridine in the axial positions, at the concentrations used for the NMR examinations which are about 4.0×10^{-3} M.

Single-crystal X-ray structure of Co-2. The **Co-2** complex was crystallized in the presence of a mixture of benzene/*n*-heptane/pyridine. The crystal was solved, and a well-defined bis(pyridine) cobalt core as shown in (Fig. 3) was obtained, exhibiting a tetragonally distorted octahedron coordination due to four practically identical nitrogen atoms of the corrole macrocycle and two nitrogen atoms of axially bounded pyridine molecules with longer metal-nitrogen bonds.

The Co–$N_{pyrrolic}$ bond lengths of the identified molecule are 1.8845 (3) Å (on average) and the Co–$N_{pyridine}$ bond distances are 2.013(3) Å which are in strong agreement with the literature value [31,38,45,48,49]. The axial bond lengths are found greater than equatorial bond distances i.e. Co-$N_{pyridine}$ > Co-$N_{pyrrolic}$ which also indicates the low spin Co(III) complex. Also these respective bond distance values are found slightly longer than the previously reported [tpFCor]Co(py)$_2$ (1.995 Å, 1.879 Å respectively) [50] and [10-(4-F-3-NO_2)Ph-5,15-di-MesitylCor]Co(py)$_2$ (1.9897 Å, 1.713825 Å respectively). [31] Table S1 display the crystallographic details of **Co-2** corrole while Table S2 compares the bond length distances to those stated in literature previously [31,50].

The cobalt ion is located at the center of a tetragonally elongated octahedron with a typical D_{4h} symmetry. The dihedral angle between the two pyridine molecules is 86.66°. The central cobalt atom **Co-2** lies in the corrole 23-atom core mean plane as well in 4-pyrrole nitrogen atoms mean plane.

UV–Visible spectroscopy for Co–(1 to 5). In the bis pyridine cobalt corroles, the pyridine molecules function as labile ligands. The **Co–(1 to 5)** complexes when prepared in the presence of pyridine, were dark green in color. However, when dried and dissolved in dichloromethane, they turn brownish green. When 1–2% of certain coordinating ligand such as pyridine was added, the solution exhibited a green color again. The appearance of the green color is more reminiscent of the 6-coordinated shape of the cobalt corrole complexes [31,51,52]. These complexes were characterized in various solvent conditions and the changes were reported using UV–Vis spectroscopy to further highlight this point.

In a 10^{-4} M DCM solution, the **Co-1** displayed major bands at 377, 450, 566, 637 nm (Fig. 4(b)) while in the 10^{-6} M DCM solution, the bands appeared at 376, 442, 566 (Fig. 4(a)). The existence of a band at around 640 nm in the more concentrated solution indicates that the complexes were somewhat present in a 6-coordinated form where the pyridine ligands occupied the axial position. While in highly diluted environments, the complex has mono-pyridine as the major form. As shown in Fig. 4, all the complexes behaved similarly. This confirms that the dilution effect has a significant impact on the coordination number of cobalt corrole complexes.

On the other hand, when the solution of these complexes was prepared in neat pyridine conditions, the coordinating solvent immediately transformed the complex into 6-coordinated form, as seen by the variations in UV–Vis spectral bands (Fig. 4(c)) and accompanied by an immediate color change from reddish-brown to green. The strong Q-band at 647 nm is the marker band for pyridine in metallocorrole that is bounded axially to cobalt metal [29,51]. The Similar band was also observed for all the **Co–(1 to 5)** complexes in 10^{-4} M DCM solution but was absent for 10^{-6} M DCM solutions. All the absorption bands and spectral variations for **Co–(1 to 5)** for DCM and pyridine solution are depicted in Fig. 4 and Table 1.

The bands associated with the UV − visible spectra of the 5- and 6-coordinate species of **Co–(1 to 5)** are close to the spectra of the mono- and bis(pyridine) adducts of other (Ar)$_3$CorCo and DPPCo derivatives previously mentioned in the literature [6,38,53]. It can therefore be inferred that the complexes are present majorly in 5-coordinated form in 10^{-6} M DCM solution, 6-coordinated in neat pyridine while present in an equilibrium condition between the 5- and 6-coordinated form at conc. 10^{-4} M DCM solution, which can be represented by (Eq. (2)).

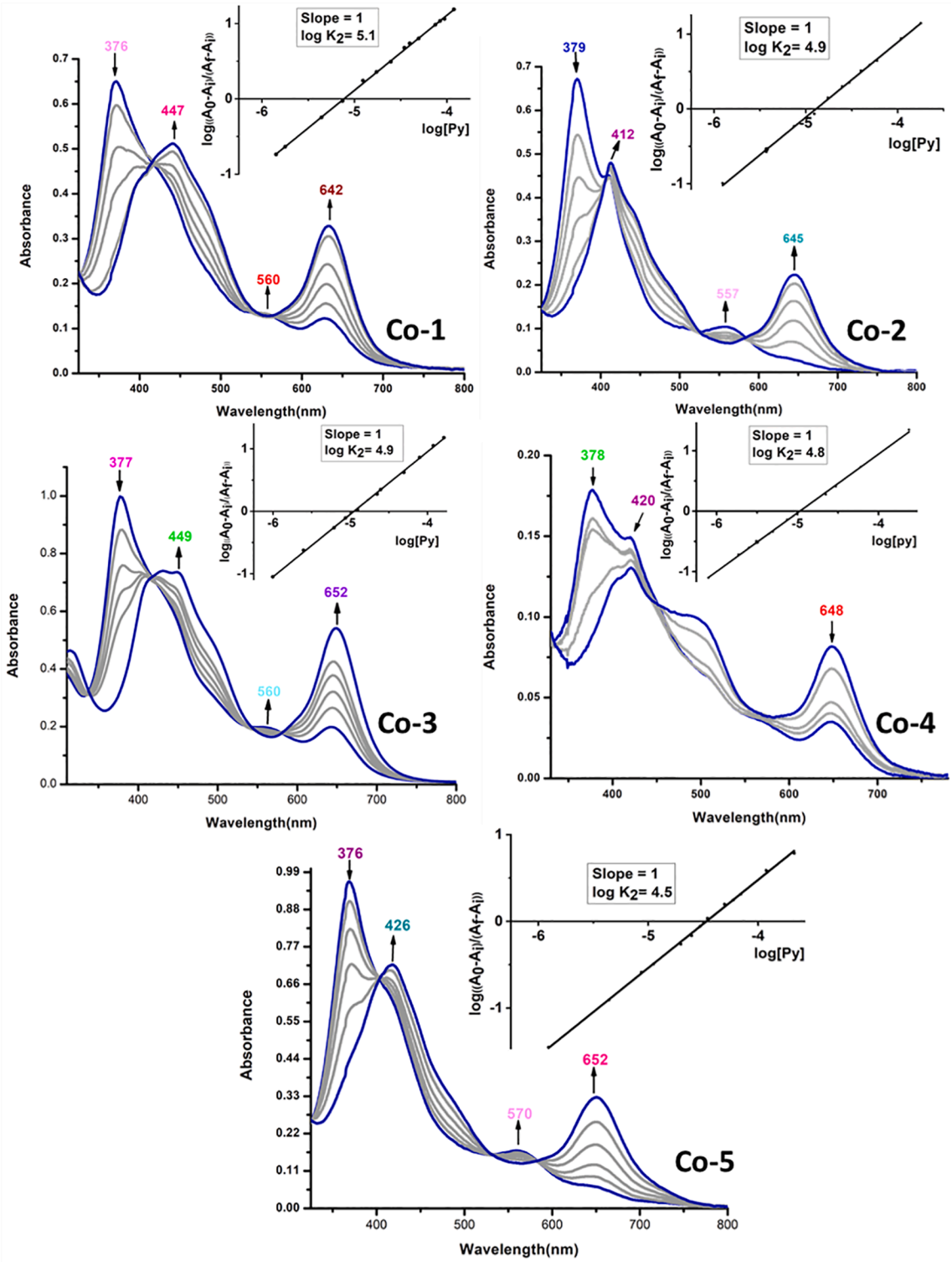


Fig. 5. Titration of **Co**–(**1 to 5**) with pyridine. Hill plots for the binding constant analysis are shown in inset.

Table 2
Pyridine binding constants for cobalt corroles in CH_2Cl_2.

Compound	Binding constant (logK)
Co-1	5.1
Co-2	4.9
Co-3	4.9
Co-4	4.8
Co-5	4.5

$$\underset{\text{Bright green}}{(Ar)_3CorCo(py)_2} \rightleftharpoons \underset{\text{Reddish}-\text{brown}}{(Ar)_3CorCo(py)} + py \quad (2)$$

The donor number is often linked to the solvent's coordination ability at axial position with cobalt-corroles [54]. As the Gutmann solvent donor number for pyridine is highest among the four below mentioned solvents (Eq. (3)), the coordinating abilities are also found higher for pyridine. Benzonitrile(PhCN) has a donor number of 11.0 as well. Hence, it could also have the capacity to bind axially with cobalt

corroles.

$$\begin{matrix} \text{DCM} < \text{PhCN} < \text{THF} < \text{Py} \\ 0.0 < 11.9 < 20.0 < 33.1 \end{matrix} \quad (3)$$

In order to analyze the benzonitrile interaction with synthesized cobalt corroles, the complexes were examined by UV–Vis spectroscopy in 10^{-5} M benzonitrile solution (Fig. S14). Similar bands for the **Co**–(**1 to 5**) complexes were observed as appearing in pyridine. For **Co-1** the bands detected were at 426, 444, and 646 nm. At 646 nm, the strong Q-band emergence, indicated that the benzonitrile solvent is functioning as a coordinating solvent in this case. Nevertheless, these effects have never been seen before in cobalt corroles. The band at 646 nm is a strong indication of the presence of a 6-coordinated form of the complexes. As our major focus was on pyridine, the rest of the study on benzonitrile behavior with cobalt corroles is beyond the scope of this paper.

To further study the coordination number of cobalt corrole complexes and to exactly determine the number of pyridine ligands binding axially, the binding constant of all the complexes was calculated which involved the titration of cobalt corrole complexes in $\sim 10^{-5}$ M DCM solution with pyridine. The titration was monitored by UV–Visible spectroscopy and at each step of pyridine addition, the UV–visible spectra of solution was recorded, and the Hill equation (Eq. (1)) was employed to calculate the logK values. The spectral changes observed during titration as well as the Hill plot for **Co**–(**1**–**5**) are depicted in (Fig. 5). A strong Q-band appearance accompanied by the reduction of intensity in Soret band was observed at the end point. The slope determined for every complex was one and the binding constant values were found to be in the 4.5–5.1 range (Table 2). The resulting slope value of one, meant that the titration consisted of one step, i.e., one pyridine ligand was bounded at axial position. This behavior is found in good accordance with the previous concentration dependent study explained above, in which the complex initially showed no spectral bands at ~ 640 nm but after the introduction of a coordinating ligand in solution, a significant behavioral shift, easily identified by UV–visible spectroscopy was found.

One key point was the binding constant values trend for the synthesized complexes i.e., the maximum logK value was observed for **Co-1** complex consisted of highly electron withdrawing group (pentafluorophenyl) at the C_{10}-*meso* position while the lowest was observed for **Co-5** onto which the group present is electron donating in nature (3,4,5-trimethoxyphenyl). This corresponds to the fact that as the electron density on the central metal decreases, the ligand binding potential enhances.

4. Conclusion

A series of A_2B type cobalt(III) corroles was prepared and studied regarding various aspects. The complexes were characterized through UV–Visible spectroscopy when present in 10^{-4} M DCM, 10^{-6} M DCM, 10^{-5} M pyridine and 10^{-5} M PhCN solution. A behavioral analysis with different solvent conditions determined the coordination number of the complexes, which is 5 in highly dilute solution or in the absence of a coordinating ligand and 6 in high concentration of DCM or in presence of coordinating ligand. Benzonitrile is also able to show 6- coordinated form, hence acting as a binding ligand. The binding constant data also revealed the binding of only one pyridine molecule to all the synthesized cobalt corrole since in very dilute solution cobalt(III) corroles exist in 5-coordinated form and hence resulting in 6-coordinated form in high concentration. These complexes are present in the 6-coordinated form in the presence of coordinating ligand such as pyridine and benzonitrile. The sharp and well resolved peaks in the ^{1}H NMR spectra proved that the complexes are present in Co(III) form. The similar results were gathered from X-ray crystallographic study of the **Co-2** complex in which the equatorial i.e., Co–$N_{pyrrolic}$ and axial i.e., Co–$N_{pyridine}$ bond distances obtained were 1.8845(3) Å (on average) and 2.013(3) Å respectively. The X-ray crystallographic and ^{1}H NMR studies hence have confirmed that bispyridine complex of cobalt triarylcorrole are authentic diamagnetic Co(III) complexes.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

AK, SGW and Jyoti thanks to Delhi Technological University for providing infrastructural facilities. Jyoti also thanks to CSIR for JRF and SRF. Authors also thanks to Professor Zeev Gross, Schulich Faculty of Chemistry, Technion-IIT, Haifa, Israel for using instrumental facilities.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ica.2021.120580.

References

[1] D.T. Gryko, J.P. Fox, D.P. Goldberg, Recent advances in the synthesis of corroles and core-modified corroles, J. Porphyr. Phthalocyanines. 08 (2002) 1091–1105, https://doi.org/10.1142/S1088424604000465.
[2] R. Guilard, J.-M. Barbe, C. Stern, K.M. Kadish, New Developments in Corrole Chemistry: Special Emphasis on Face-to-Face Bismacrocycles, in, The Porphyrin Handbook, Elsevier (2003) 303–349, https://doi.org/10.1016/B978-0-08-092392-5.50011-0.
[3] R. Paolesse, Corrole: The little big porphyrinoid, Synlett. 2008 (15) (2008) 2215–2230, https://doi.org/10.1055/s-2008-1078687.
[4] J.H. Palmer, Transition Metal Corrole Coordination Chemistry, in, Springer-Verlag, Berlin Heidelberg, 2011, pp. 49–90.
[5] S. Nardis, F. Mandoj, M. Stefanelli, R. Paolesse, F. Sara, M. Mandoj, R.P. Stefanelli, Metal complexes of corrole, Coord. Chem. Rev. 388 (2019) 360–405, https://doi.org/10.1016/j.ccr.2019.02.034.
[6] V.A. Adamian, F. D'Souza, S. Licoccia, M.L. Di Vona, E. Tassoni, R. Paolesse, T. Boschi, K.M. Kadish, Synthesis, Characterization, and Electrochemical Behavior of (5,10,15-Tri-X-phenyl-2,3,7,8,12,13,17,18-octamethylcorrolato)cobalt(III) Triphenylphosphine Complexes, Where X = p-OCH_3, p-CH_3, p-Cl, m-Cl, o-Cl, m-F, o-F, or H, Inorg. Chem. 34 (1995) 532–540, https://doi.org/10.1021/ic00107a003.
[7] K.M. Kadish, J. Shen, L. Frémond, P. Chen, M. El Ojaimi, M. Chkounda, C.P. Gros, J.-M. Barbe, K. Ohkubo, S. Fukuzumi, R. Guilard, Clarification of the oxidation state of cobalt corroles in heterogeneous and homogeneous catalytic reduction of dioxygen, Inorg. Chem. 47 (15) (2008) 6726–6737, https://doi.org/10.1021/ic800458s.
[8] A. Ghosh, Electronic Structure of Corrole Derivatives: Insights from Molecular Structures, Spectroscopy, Electrochemistry, and Quantum Chemical Calculations, Chem. Rev. 117 (4) (2017) 3798–3881, https://doi.org/10.1021/acs.chemrev.6b00590.
[9] K.M. Kadish, L. Frémond, Z. Ou, J. Shao, C. Shi, F.C. Anson, F. Burdet, C.P. Gros, J. M. Barbe, R. Guilard, Cobalt(III) corroles as electrocatalysts for the reduction of dioxygen: Reactivity of a monocorrole, biscorroles, and porphyrin-corrole dyads, J. Am. Chem. Soc. 127 (2005) 5625–5631, https://doi.org/10.1021/ja0501060.
[10] J.P. Collman, M. Kaplun, R.A. Decréau, Metal corroles as electrocatalysts for oxygen reduction, Dalt. Trans. (4) (2006) 554–559, https://doi.org/10.1039/B512982F.
[11] Z. Ou, A. Lü, D. Meng, S. Huang, Y. Fang, G. Lu, K.M. Kadish, Molecular Oxygen Reduction Electrocatalyzed by meso -Substituted Cobalt Corroles Coated on Edge-Plane Pyrolytic Graphite Electrodes in Acidic Media, Inorg. Chem. 51 (16) (2012) 8890–8896, https://doi.org/10.1021/ic300886s.
[12] A. Mahammed, Z. Gross, Metallocorroles as Electrocatalysts for the Oxygen Reduction Reaction (ORR), Isr. J. Chem. 56 (9-10) (2016) 756–762, https://doi.org/10.1002/ijch.v56.9-1010.1002/ijch.201600027.
[13] H. Sun, Y. Han, H. Lei, M. Chen, R. Cao, Cobalt corroles with phosphonic acid pendants as catalysts for oxygen and hydrogen evolution from neutral aqueous solution, Chem. Commun. 53 (46) (2017) 6195–6198, https://doi.org/10.1039/C7CC02400B.
[14] X. Li, H. Lei, J. Liu, X. Zhao, S. Ding, Z. Zhang, X. Tao, W. Zhang, W. Wang, X. Zheng, R. Cao, Carbon Nanotubes with Cobalt Corroles for Hydrogen and Oxygen Evolution in pH 0–14 Solutions, Angew. Chemie - Int. Ed. 57 (2018) 15070–15075, https://doi.org/10.1002/anie.201807996.
[15] A. Kumar, S. Sujesh, P. Varshney, A. Paul, S. Jeyaraman, Aminophenyl-substituted cobalt(III) corrole: A bifunctional electrocatalyst for the oxygen and hydrogen evolution reactions, Dalt. Trans. 48 (2019) 11345–11351, https://doi.org/10.1039/c9dt02339a.

[16] D. Dolui, S. Ghorai, A. Dutta, Tuning the reactivity of cobalt-based H_2 production electrocatalysts via the incorporation of the peripheral basic functionalities, Coord. Chem. Rev. 416 (2020) 213335, https://doi.org/10.1016/j.ccr.2020.213335.

[17] H. Lin, M.S. Hossain, S.Z. Zhan, H.Y. Liu, L.P. Si, Electrocatalytic hydrogen evolution using triaryl corrole cobalt complex, Appl. Organomet. Chem. (2020) 1–10, https://doi.org/10.1002/aoc.5583.

[18] H. Lei, X. Li, J. Meng, H. Zheng, W. Zhang, R. Cao, Structure Effects of Metal Corroles on Energy-Related Small Molecule Activation Reactions, ACS Catal. 9 (5) (2019) 4320–4344, https://doi.org/10.1021/acscatal.9b00310.

[19] J. Grodkowski, P. Neta, E. Fujita, A. Mahammed, L. Simkhovich, Z. Gross, Reduction of cobalt and iron corroles and catalyzed reduction of CO_2, J. Phys. Chem. A. 106 (18) (2002) 4772–4778, https://doi.org/10.1021/jp013668o.

[20] W. Sinha, A. Mahammed, N. Fridman, Y. Diskin-Posner, L.J.W. Shimon, Z. Gross, Superstructured metallocorroles for electrochemical CO_2 reduction, Chem. Commun. 55 (79) (2019) 11912–11915, https://doi.org/10.1039/C9CC06645D.

[21] S. Gonglach, S. Paul, M. Haas, F. Pillwein, S.S. Sreejith, S. Barman, R. De, S. Müllegger, P. Gerschel, U.P. Apfel, H. Coskun, A. Aljabour, P. Stadler, W. Schöfberger, S. Roy, Molecular cobalt corrole complex for the heterogeneous electrocatalytic reduction of carbon dioxide, Nat. Commun. 10 (2019) 1–10, https://doi.org/10.1038/s41467-019-11868-5.

[22] S. Brandès, V. Quesneau, O. Fonquernie, N. Desbois, V. Blondeau-Patissier, C. P. Gros, Porous organic polymers based on cobalt corroles for carbon monoxide binding, Dalt. Trans. 48 (31) (2019) 11651–11662, https://doi.org/10.1039/C9DT01599J.

[23] J.-M. Barbe, G. Canard, S. Brandès, R. Guilard, Synthesis and physicochemical characterization of meso-functionalized corroles: Precursors of organic-inorganic hybrid materials, European, J. Org. Chem. 2005 (21) (2005) 4601–4611, https://doi.org/10.1002/(ISSN)1099-069010.1002/ejoc.v2005:2110.1002/ejoc.200500374.

[24] Y. Wang, J. Akhigbe, Y.u. Ding, C. Brückner, Y.u. Lei, Meso-Tritolylcorrole-Functionalized Single-walled Carbon Nanotube Donor-Acceptor Nanocomposites for NO_2 Detection, Electroanalysis. 24 (6) (2012) 1348–1355, https://doi.org/10.1002/elan.v24.610.1002/elan.201200077.

[25] S.i. Yang, M.E. Meyerhoff, Study of cobalt(III) corrole as the neutral ionophore for nitrite and nitrate detection via polymeric membrane electrodes, Electroanalysis. 25 (12) (2013) 2579–2585, https://doi.org/10.1002/elan.v25.1210.1002/elan.201300400.

[26] X. Sheng, H. Zhao, L. Du, Selectivity of Cobalt Corrole for CO vs. O_2 and N_2 in Indoor Pollution, Sci. Rep. 7 (2017) 14536, https://doi.org/10.1038/s41598-017-15228-5.

[27] E. Jaworska, G. Pomarico, B.B. Berna, K. Maksymiuk, R. Paolesse, A. Michalska, All-solid-state paper based potentiometric potassium sensors containing cobalt(II) porphyrin/cobalt(III) corrole in the transducer layer, Sensors Actuators, B Chem. 277 (2018) 306–311, https://doi.org/10.1016/j.snb.2018.08.090.

[28] M. Vanotti, S. Poisson, V. Soumann, V. Quesneau, S. Brandès, N. Desbois, J. Yang, L. André, C.P. Gros, V. Blondeau-Patissier, Influence of interfering gases on a carbon monoxide differential sensor based on SAW devices functionalized with cobalt and copper corroles, Sensors Actuators, B Chem. 332 (2021) 129507, https://doi.org/10.1016/j.snb.2021.129507.

[29] V. Quesneau, W. Shan, N. Desbois, S. Brandès, Y. Rousselin, M. Vanotti, V. Blondeau-Patissier, M. Naitana, P. Fleurat-Lessard, E. Van Caemelbecke, K. M. Kadish, C.P. Gros, Cobalt Corroles with Bis-Ammonia or Mono-DMSO Axial Ligands. Electrochemical, Spectroscopic Characterizations and Ligand Binding Properties, Eur. J. Inorg. Chem. 2018 (38) (2018) 4265–4277, https://doi.org/10.1002/ejic.201800897.

[30] X. Jiang, W. Shan, N. Desbois, V. Quesneau, S. Brandès, E.V. Caemelbecke, W. R. Osterloh, V. Blondeau-Patissier, C.P. Gros, K.M. Kadish, Mono-DMSO ligated cobalt nitrophenylcorroles: Electrochemical and spectral characterization, New J. Chem. 42 (10) (2018) 8220–8229, https://doi.org/10.1039/C8NJ00300A.

[31] X. Jiang, M.L. Naitana, N. Desbois, V. Quesneau, S. Brandès, Y. Rousselin, W. Shan, W.R. Osterloh, V. Blondeau-Patissier, C.P. Gros, K.M. Kadish, Electrochemistry of Bis(pyridine)cobalt (Nitrophenyl)corroles in Nonaqueous Media, Inorg. Chem. 57 (3) (2018) 1226–1241, https://doi.org/10.1021/acs.inorgchem.7b0265510.1021/acs.inorgchem.7b02655.s001.

[32] W. Sinha, A. Mizrahi, A. Mahammed, B. Tumanskii, Z. Gross, Reactive Intermediates Involved in Cobalt Corrole Catalyzed Water Oxidation (and Oxygen Reduction), Inorg. Chem. 57 (1) (2018) 478–485, https://doi.org/10.1021/acs.inorgchem.7b0269610.1021/acs.inorgchem.7b02696.s001.

[33] K.M. Kadish, J. Shao, Z. Ou, L. Frémond, R. Zhan, F. Burdet, J.-M. Barbe, C.P. Gros, R. Guilard, Electrochemistry, spectroelectrochemistry, chloride binding, and O_2 catalytic reactions of free-base porphyrin - Cobalt corrole dyads, Inorg. Chem. 44 (19) (2005) 6744–6754, https://doi.org/10.1021/ic050738n10.1021/ic050738n.s001.

[34] W. Lee, X. Zhan, J. Palma, J. Vestfrid, Z. Gross, D.G. Churchill, Minding our P-block and Q-bands: Paving inroads into main group corrole research to help instil broader potential, Chem. Commun. 57 (38) (2021) 4605–4641, https://doi.org/10.1039/D1CC00105A.

[35] W.R. Osterloh, V. Quesneau, N. Desbois, S. Brandès, W. Shan, V. Blondeau-Patissier, R. Paolesse, C.P. Gros, K.M. Kadish, Synthesis and the Effect of Anions on the Spectroscopy and Electrochemistry of Mono(dimethyl sulfoxide)-Ligated Cobalt Corroles, Inorg. Chem. 59 (1) (2020) 595–611, https://doi.org/10.1021/acs.inorgchem.9b0285510.1021/acs.inorgchem.9b02855.s001.

[36] W.R. Osterloh, N. Desbois, V. Quesneau, S. Brandès, P. Fleurat-Lessard, Y. Fang, V. Blondeau-Patissier, R. Paolesse, C.P. Gros, K.M. Kadish, Old Dog, New Tricks: Innocent, Five-coordinate Cyanocobalt Corroles, Inorg. Chem. 59 (12) (2020) 8562–8579, https://doi.org/10.1021/acs.inorgchem.0c0103710.1021/acs.inorgchem.0c01037.s001.

[37] P.E. Ellis, J.E. Linard, T. Szymanski, R.D. Jones, J.R. Budge, F. Basolo, Axial Ligation Constants of Iron(II) and Cobalt(II) "Capped" Porphyrins, J. Am. Chem. Soc. 102 (6) (1980) 1889–1896, https://doi.org/10.1021/ja00526a022.

[38] R. Guilard, C.P. Gros, F. Bolze, F. Jérôme, Z. Ou, J. Shao, J. Fischer, R. Weiss, K. M. Kadish, Alkyl and aryl substituted corroles. 1. Synthesis and characterization of free base and cobalt containing derivatives. X-ray structure of $(Me_4Ph_5Cor)Co(py)_2$, Inorg. Chem. 40 (2001) 4845–4855, https://doi.org/10.1021/ic010177+.

[39] R. Guilard, F. Jérôme, J.-M. Barbe, C.P. Gros, Z. Ou, J. Shao, J. Fischer, R. Weiss, K. M. Kadish, Alkyl and aryl substituted corroles. 2. Synthesis and characterization of linked "face-to-face" biscorroles. X-ray structure of $(BCA)Co_2(py)_3$, where BCA represents a biscorrole with an anthracenyl bridge, Inorg. Chem. 40 (19) (2001) 4856–4865, https://doi.org/10.1021/ic0101782.

[40] B. Koszarna, D.T. Gryko, Efficient synthesis of meso-substituted corroles in a H_2O-MeOH mixture, J. Org. Chem. 71 (10) (2006) 3707–3717, https://doi.org/10.1021/jo060007k10.1021/jo060007k.s00110.1021/jo060007k.s002.

[41] A. Mahammed, I. Giladi, I. Goldberg, Z. Gross, Synthesis and Structural Characterization of a novel Covalently-Bound Corrole Dimer, Chem. Eur. J. 2001,. 723 (2001) 4259–4265. https://doi.org/0947-6539/01/0719-4259 $ 17.50+.50/0.

[42] O.V. Dolomanov, L.J. Bourhis, R.J. Gildea, J.A.K. Howard, H. Puschmann, *OLEX2*: A complete structure solution, refinement and analysis program, J. Appl. Crystallogr. 42 (2) (2009) 339–341, https://doi.org/10.1107/S0021889808042726.

[43] L.J. Bourhis, O.V. Dolomanov, R.J. Gildea, J.A.K. Howard, H. Puschmann, The anatomy of a comprehensive constrained, restrained refinement program for the modern computing environment - *Olex2* dissected, Acta Crystallogr. Sect. A Found. Crystallogr. 71 (2015) 59–75, https://doi.org/10.1107/S2053273314022207.

[44] G.M. Sheldrick, Crystal structure refinement with *SHELXL*, Acta Crystallogr. Sect. C Struct. Chem. 71 (2015) 3–8, https://doi.org/10.1107/S2053229614024218.

[45] S. Ganguly, J. Conradie, J. Bendix, K.J. Gagnon, L.J. McCormick, A. Ghosh, Electronic Structure of Cobalt–Corrole–Pyridine Complexes: Noninnocent Five-Coordinate Co(II) Corrole-Radical States, J. Phys. Chem. A. 121 (50) (2017) 9589–9598, https://doi.org/10.1021/acs.jpca.7b0944010.1021/acs.jpca.7b09440.s00110.1021/acs.jpca.7b09440.s002.

[46] L. Simkhovich, N. Galili, I. Saltsman, I. Goldberg, Z. Gross, Coordination chemistry of the novel 5,10,15-tris(pentafluorophenyl)corrole: Synthesis, spectroscopy, and structural characterization of its cobalt(III), rhodium(III), and iron(IV) complexes, Inorg. Chem. 39 (2000) 2704–2705, https://doi.org/10.1021/ic991342c.

[47] K. Sudhakar, A. Mahammed, N. Fridman, Z. Gross, Trifluoromethylation for affecting the structural, electronic and redox properties of cobalt corroles, Dalt. Trans. 48 (15) (2019) 4798–4810, https://doi.org/10.1039/C9DT00675C.

[48] S. Ooi, T. Tanaka, A. Osuka, Cobalt(III) and gallium(III) complexes of meso-free corroles with distinct position-dependent substituent effects, J. Porphyr. Phthalocyanines. 20 (01n04) (2016) 274–281, https://doi.org/10.1142/S1088424616500061.

[49] B. Mondal, K. Sengupta, A. Rana, A. Mahammed, M. Botoshansky, S.G. Dey, Z. Gross, A. Dey, Cobalt Corrole Catalyst for Efficient Hydrogen Evolution Reaction from H_2O under Ambient Conditions: Reactivity, Spectroscopy, and Density Functional Theory Calculations, Inorg. Chem. 52 (6) (2013) 3381–3387, https://doi.org/10.1021/ic4000473.

[50] H. Lei, A. Han, F. Li, M. Zhang, Y. Han, P. Du, W. Lai, R. Cao, Electrochemical, spectroscopic and theoretical studies of a simple bifunctional cobalt corrole catalyst for oxygen evolution and hydrogen production, Phys. Chem. Chem. Phys. 16 (5) (2014) 1883–1893, https://doi.org/10.1039/C3CP54361G.

[51] K.M. Kadish, Z. Ou, J. Shao, C.P. Gros, J.M. Barbe, F. Jérôme, F. Bolze, F. Burdet, R. Guilard, Alkyl and aryl substituted corroles. 3. Reactions of cofacial cobalt biscorroles and porphyrin-corroles with pyridine and carbon monoxide, Inorg. Chem. 41 (2002) 3990–4005, https://doi.org/10.1021/ic020058+.

[52] S. Ganguly, D. Renz, L.J. Giles, K.J. Gagnon, L.J. McCormick, J. Conradie, R. Sarangi, A. Ghosh, Cobalt- and Rhodium-Corrole-Triphenylphosphine Complexes Revisited: The Question of a Noninnocent Corrole, Inorg. Chem. 56 (24) (2017) 14788–14800, https://doi.org/10.1021/acs.inorgchem.7b0182810.1021/acs.inorgchem.7b01828.s001.

[53] W. Shan, N. Desbois, S. Pacquelet, Stéphane Brandès, Y. Rousselin, J. Conradie, A. Ghosh, C.P. Gros, K.M. Kadish, Ligand Noninnocence in Cobalt Dipyrrin-Bisphenols: Spectroscopic, Electrochemical, and Theoretical Insights Indicating an Emerging Analogy with Corroles, Inorg. Chem. 58 (12) (2019) 7677–7689, https://doi.org/10.1021/acs.inorgchem.8b0300610.1021/acs.inorgchem.8b03006.s001.

[54] K.M. Kadish, P.W. Koh, Electrochemistry of Rhodium and Cobalt Corroles. Characterization of (OMC)Rh(PPh3) and (OMC)Co(PPh3) Where OMC is the Trianion of 2,3,7,8,12,13,17,18-Octamethylcorrole, Am. Chem. Soc. 31 (1992) 2305–2313, https://doi.org/0020-1669/92/1331-2305$03.00/0.

Eur. Phys. J. Plus (2021) 136:820
https://doi.org/10.1140/epjp/s13360-021-01821-1

THE EUROPEAN PHYSICAL JOURNAL PLUS

Regular Article

The generalized second law of thermodynamics in viscous Ricci dark energy model

Ajay Kumar [a], **C. P. Singh**[b]

Department of Applied Mathematics, Delhi Technological University, Bawana Road, Delhi 110 042, India

Received: 8 December 2020 / Accepted: 30 July 2021
© The Author(s), under exclusive licence to Società Italiana di Fisica and Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract In this paper, we study the validity of the generalized second law of thermodynamics by applying Bekestein–Hawking (BH) entropy and Barrow entropy for the horizon entropy, respectively in the viscous holographic Ricci dark energy (HRDE) model. The background is a spatially flat Friedmann–Robertson–Walker universe filled with pressureless dark matter and viscous HRDE. The bulk viscous coefficient is assumed to be $\zeta = \zeta_0 + \zeta_1 H$. In first step, we find the analytical solution of the field equations of viscous HRDE model. We use the observational data from SNe, OHD and BAO/CMB in order to extract the constraints on viscous HRDE model. In second step, we calculate the entropy time-variation for each fluid component and for the apparent horizon itself. In BH entropy case, the sum of the entropy of the fluids (viscous HRDE and dark matter) enclosed by the apparent horizon plus the entropy of the horizon itself is non-decreasing function of time which shows that the generalized second law of thermodynamics is valid. It also satisfies the convexity condition which shows that the model approaches a stable equilibrium thermal state. However, in the case of Barrow entropy, the generalized second law of thermodynamics may be violated as it depends on the universe evolution. Finally, we examine the generalized second law of thermodynamics of viscous HRDE model by taking into account the Casimir effect and find that it is valid throughout the evolution of the universe.

1 Introduction

A considerable number of high quality of cosmological observations [1–6] show that the present universe is accelerating and its geometry is spatially flat. The analysis of the observational data shows that the universe consist of nearly three-forth hypothetic energy component with negative pressure called 'dark energy' (DE), which is responsible for this late-time acceleration. The remaining content of the universe is so-called dark matter (DM), which is responsible for the formation of the galaxies, and the baryonic matter [7–9]. The easiest way to fit the current cosmological data is to include cosmological constant (CC) in the Friedmann equations. However, this model, so-called concordance model (or ΛCDM model) suffers from the CC problem [10]. In order to address CC problems, dynamical DE models

[a] e-mail: maths.ajy@gmail.com

[b] e-mail: cpsphd@rediffmail.com (corresponding author)

Published online: 09 August 2021

such as quintessence [11–13], K-essence [14,15], tachyon [16,17], phantom [18], Chaplygin gas model [19], holographic DE [20,21], agegraphic DE [22,23], etc. have been proposed to explain the accelerating expansion of the universe.

In recent years, to understand the nature of the universe, a new dark energy model has been proposed in the context of quantum gravity which is based on holographic principle (HP) and is known as holographic dark energy (HDE). Holographic principle is an important result of the recent research for exploring the quantum gravity (or string theory). When considering gravity in a quantum gravity system, the conventional local quantum field theory will break down because it contains too many degrees of freedom that will lead to the formation of a black hole breaking the effectiveness of the quantum field theory. It is realized that in quantum gravity, the entropy of a system scales not with its volume, but with its surface area L^2.

t'Hooft [24] and Susskind [25] used the laws of black hole thermodynamics to argue for a general holographic principle which asserts that consistent theories of gravity and quantum mechanics must be lower-dimension [26]. Black hole thermodynamics tells us that a black hole has an entropy proportional to its horizon area. In Einstein gravity theory, black hole entropy satisfies the so-called formula, $S = A/4G$, where A is the horizon area of the black hole and G is the Newtonian constant. Following HP, one may put an energy bound on the vacuum energy density $L^3\rho_{de} \leq LM_P^2$, where M_P^2 is the reduced Planck mass. This inequality supports the study done by Karolyhazy [27]. In this work, the quantum mechanics was combined with general relativity. Actually, Heisenberg uncertainty relation and gravity are combined and Karolyhazy derived a quantitative limitation on the uncertainty of the structure of space-time. Almost all HDE models are based on this pioneering combination.

It is to be noted that in Einstein equation $G_{\mu\nu} = 8\pi G T_{\mu\nu} + \Lambda g_{\mu\nu}$, the cosmological constant Λ is the inverse of some length squared, $\Lambda \sim L^{-2}$ and to be consistent with the observations, L must be of same order as the present cosmological scale. Since the cosmological constant is entangled with the vacuum expectation value of some quantum fields, the cosmological constant problem therefore is essentially an issue of quantum gravity. Indeed, general relativity together with quantum field theory could shed some lights on the cosmological constant problem. To realize this arguments, Cohen et al. [28] proposed that an unknown vacuum energy could be present, and according to the holographic principle its density is proportional to the Hubble scale $L_H \sim H^{-1}$. In this context, Li [21] proposed a model to interpret the positive acceleration of the universe with dark energy density $\rho_{de} \propto L^{-2}$, where L represents infra-red (IR) length scale cutoff of the model. It is to be noted that the HDE model depends on the choices of the IR cutoff L. In recent years, different types of HDE models have been proposed which depend on the choices of the various IR cutoff.

Inspired by HDE models, Gao et al. [29] proposed another possibility of IR cut-off where the length L is given by the average radius of Ricci scalar curvature, $|R|^{-1/2}$, so that the dark energy density has the form $\rho_{de} \propto R$, known as "holographic Ricci dark energy" (HRDE) model. The Ricci scalar curvature in the flat Friedmann–Robertson–Walker universe is given by $R = -6(\dot{H} + 2H^2)$, where H is the Hubble parameter and the dot denotes the derivative with respect to the cosmic time t. Thus, the DE density of HRDE in a spatially flat universe can be expressed as [29]

$$\rho_{de} = 3\alpha(\dot{H} + 2H^2), \tag{1}$$

where α is a constant to be determined by observations.

The studies show that this model works fairly well in explaining the observations such as the cosmic acceleration, and it could also help to understand the cosmic coincidence problem. Many authors [30–46] have studied the HRDE model to explain the accelerated expansion of the universe.

On the other hand, dissipative effect (which includes both bulk and shear viscosity) plays a very important role in the evolution of the universe. Eckart [47] was the first one to introduce a theory of relativistic dissipative fluids, and later on Landau and Lifshitz [48] modified this theory. A relativistic second-order theory was developed by Israel and Stewart [49], known as full causal theory. However, the second-order theory is so complicated that the simplified phenomenon, i.e., the quasi-stationary phenomenon, is studied by the conventional theory. In isotropic and homogeneous universe, the dissipative process is usually characterized by a bulk viscosity. The effect of the shear viscosity is neglected as the CMBR observations indicate an isotropic universe [50].

According to thermodynamics, when the local thermodynamics equilibrium is broken, the bulk viscosity arises [51]. The bulk viscosity can be treated as an effective pressure which keeps the system back to its thermal equilibrium state. In a cosmological fluid, the bulk viscosity arises when the fluid expands (or contract) too fast so that the system does not have enough time to restore the local thermodynamic equilibrium. This creates an effective pressure to restore the system to its local thermodynamic equilibrium. When the fluid reaches again the thermal equilibrium, the bulk viscosity pressure ceases [52]. It is therefore natural to assume the possibility of bulk viscosity in an accelerating expanding universe since expansion process is likely to be a collection of states out of thermal equilibrium in a small fraction of time. Thus, there exists a bulk viscosity in the description of the universe. In the papers [53,54], the problems associated with viscous cosmology have been discussed in details. In the late accelerating universe, the effect of bulk viscosity on the cosmological models have been studied by many authors [50,55–72]. Some authors have studied the viscous model with cosmological constant (negative) to describe the thermodynamics and phase transition of the early evolution of the universe [73–80].

There is well-known connection between black-hole physics and thermodynamics. In 1973, Motivated by the similarities between the physics of black-hole and thermodynamics, Bekenstein [81–83] demonstrated that the entropy of black-hole is proportional to its event horizon. His work was also supported by Hawking [84], which leads to an important generalized second law (GSL) of thermodynamics for black-hole physics. The temperature and entropy of black-hole depend on the black-hole horizon [85]. Thermodynamics tells us that the entropy of isolated systems can never diminishes. Cai et al. [86] demonstrated that cosmological apparent horizons are endowed with thermodynamical properties (also, refer to Ref. [87]) One can relate the temperature and entropy to the apparent horizon analogous to ones associated to the black-hole event horizon. The entropy of the apparent horizon is given by $S_h = \kappa_B A/4l_{Pl}^2$, where $A = 4\pi\tilde{r}_A^2$ is the area of the horizon, κ_B the Boltzmann's constant, l_{Pl} the Planck's length, and $\tilde{r}_A$ the radius of the horizon [88]. Recently, Barrow [89] has analyzed that quantum gravitational effects could bring about intricate, fractal structure on the black-hole surface and thus changes its actual horizon area which leads to a new black-hole entropy relation $S_B = (A/A_0)^{(1+\Delta/2)}$, where A_0 is the Planck area. The quantum gravitational perturbation is represented by the new exponent Δ. According to the GSL the entropy of matter and fields inside the horizon plus the entropy of the horizon cannot decrease with time. Recent study on GSL of thermodynamics in various gravity theories can be found in the literature [90–102].

It is to be noted that in papers [67,68] the authors presented a theoretical study of HRDE with bulk viscosity with some aspects of thermodynamics. We used previous data obtained in [38] and various combinations of other model parameters to discuss the evolution of the universe. In this present paper, we extend these study to the viscous model via a qualitatively numerical analysis. We place the constraints on model parameters by performing a Bayesian analysis based on Markov Chain Monte Carlo method and using the latest data of

the Hubble parameter (OHD), Type Ia supernovae (SNe) and Baryonic acoustic oscillation (BAO)/Cosmic microwave background (CMB). Using the best-fit values, we examine the validity of the GSL of thermodynamics with different black-hole entropy like, Bekestein–Hawking, Barrow entropy and Casimir effect on the apparent horizon.

The paper is organized as follows: In Sect. 2, we present the viscous HRDE model and its analytic solution. Section 3 deals with the parameter estimation using observational data from SNe, OHD and BAO/CMB to constraint the viscous HRDE model. In Sect. 4, we present the generalized second law of thermodynamics with Bekestein–Hawking entropy for viscous HRDE model by considering the universe as a system bounded by the apparent horizon. In Sect. 5, we discuss the generalized second law of thermodynamics with Barrow entropy which arises from the modification of the black-hole surface due to quantum gravitational effects. In Sect. 6, we extend our study by taking Casimir effect into consideration. Finally, Sect. 7 is devoted to the summary of the obtained results.

2 Viscous HRDE model

We consider a spatially homogeneous and isotropic flat universe described by the Friedmann–Robertson–Walker (FRW) metric

$$\mathrm{d}s^2 = dt^2 - a^2(t)\left[dr^2 + r^2(d\theta^2 + \sin^2\theta \mathrm{d}\phi^2)\right], \tag{2}$$

where $a(t)$ represents the cosmic scale factor of the universe. For simplicity, throughout this work we are using units where $8\pi G = c = 1$.

We consider that the universe is filled with pressureless dark matter (excluding the baryonic matter) and viscous HRDE for which the stress-energy-momentum tensor is given as

$$T_{\mu\nu} = (\rho_m + \rho_{de} + \tilde{p_{de}})u_\mu u_\nu - \tilde{p_{de}}\, g_{\mu\nu}, \tag{3}$$

where ρ_m and ρ_{de} are, respectively, the energy density of dark matter and HRDE, and $\tilde{p_{de}}$ represents the effective pressure of HRDE and viscous fluid, i.e., $\tilde{p_{de}} = p_{de} - 3\zeta H$, where p_{de} is the pressure of HRDE, ζ is the bulk viscous coefficient and $H = \dot{a}/a$ is the scalar expansion. The condition $\zeta > 0$ guaranties a positive entropy production and, in consequence, no violation of the second law of thermodynamics [103].

The bulk viscous coefficient depends upon the cosmic time t, or the scale factor a, or redshift z, or their combinations. Wang et al. [66] assumed that the bulk viscous coefficient is proportional to the Hubble parameter, i.e., $\zeta = \zeta_0 H$ and presented the viscous dark energy model with cosmological constant. In general, this form of bulk viscous coefficient gives power-law expansion of the universe, i.e., the universe decelerates or accelerates depending on the power values of the scale factor. This also gives constant deceleration parameter which does not show the transition. This type of result has been found in case (i) of Sect. 3.2 in our paper [67], which is not suitable in this current scenario of accelerating universe. Therefore, in our present model, we assume a more generalized form of bulk viscous coefficient by adding a constant term [63],

$$\zeta = \zeta_0 + \zeta_1 H, \tag{4}$$

where ζ_0 and ζ_1 are positive constants. Now, we define the dimensionless bulk viscous parameters ξ, ξ_0 and ξ_1 as

$$\xi = \frac{\zeta}{H_0}, \quad \xi_0 = \frac{\zeta_0}{H_0}, \quad \xi_1 = \zeta_1, \tag{5}$$

where H_0 is the present value of the Hubble parameter. Using (5) into (4), we obtain the dimensionless form of bulk viscous parameter as

$$\xi = \xi_0 + \xi_1 h, \tag{6}$$

where $h = \frac{H}{H_0}$ is the dimensionless Hubble parameter.

The Friedmann equation for a flat universe (2) filled with energy-momentum tensor (3) takes the form

$$3H^2 = \rho_m + \rho_{de}. \tag{7}$$

We assume that dark matter and viscous HRDE do not interact with each other, and therefore, the energy conservation equation of both the fluids conserve separately, which are given by

$$\dot{\rho}_m + 3H\rho_m = 0, \tag{8}$$

and

$$\dot{\rho}_{de} + 3H\rho_{de}(1 + \omega_{de}) = 9\zeta H^2, \tag{9}$$

where $\omega_{de} = p_{de}/\rho_{de}$ is the equation of state (EoS) parameter of HRDE. Using (1) and (6) into (7)–(9), we obtain a single dimensionless evolution equation

$$(2 + 3\alpha\omega_{de})h' + 3(1 - \xi_1 + 2\alpha\omega_{de})h - 3\xi_0 = 0, \tag{10}$$

where a prime denotes derivative with respect to conformal time $ln\ a$. On solving (10), we obtain

$$h(a) = \frac{\xi_0}{(1 - \xi_1 + 2\alpha\omega_{de})} + \left(1 - \frac{\xi_0}{(1 - \xi_1 + 2\alpha\omega_{de})}\right)\left(\frac{a}{a_0}\right)^{-\frac{3(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}} \tag{11}$$

Integrating (11), we obtain the scale factor as

$$a(t) = a_0\left[1 + \frac{(1 - \xi_1 + 2\alpha\omega_{de})}{\xi_0}\left(e^{\frac{3\xi_0 H_0(t-t_0)}{(2+3\alpha\omega_{de})}} - 1\right)\right]^{\frac{(2+3\alpha\omega_{de})}{3(1-\xi_1+2\alpha\omega_{de})}} \tag{12}$$

where a_0 is the current value of the scale factor at time $t = t_0$, and $\xi_0 \neq 0$ and $\xi_1 \neq (1 + 2\alpha\omega_{de})$. Now onwards, we assume $a_0 = 1$. It is observed that in the early time, the scale factor varies as power-law form, i.e., $a \propto t^m$, $(m \geq 0)$, whereas in the late-time, the scale factor varies as $a \propto \exp[3\xi_0 H_0(t - t_0)/(2 + 3\alpha\omega_{de})]$, which shows the de Sitter phase of the universe.

In cosmology, the deceleration parameter q plays an important role in describing the nature of the expansion of the universe. It is defined as $q = -a\ddot{a}/\dot{a}^2$. For viscous HRDE model, q is given by

$$q = -1 + \frac{3(1 - \xi_0 - \xi_1 + 2\alpha\omega_{de})(1 + z)^{\frac{3(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}}}{(2 + 3\alpha\omega_{de})\left[\frac{\xi_0}{(1-\xi_1+2\alpha\omega_{de})} + \left(1 - \frac{\xi_0}{(1-\xi_1+2\alpha\omega_{de})}\right)(1 + z)^{\frac{3(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}}\right]}, \tag{13}$$

where $1 + z = a_0/a$ is the redshift.

Table 1 Best fit results for viscous HRDE model using joint analysis of data set from SNe, BAO/CMB and OHD

Parameters	Best fit value
ξ_0	$0.603^{+0.112}_{-0.164}$
ξ_1	$-0.149^{+0.031}_{-0.047}$
α	$0.726^{+1.049}_{-0.389}$
ω_{de}	$-0.136^{+0.092}_{-0.174}$

3 Parameters estimation

To analyze the viscous HRDE model and to obtain the best fit values for the model parameters, we combine the latest observational data including SNe, OHD and BAO/CMB (see, "Appendix A" for detail).

- *Supernova Type Ia (SNe):* SNe data were the first to indicate that our universe is accelerating. In this paper, we include the latest Joint Light Curve analysis (JLA) sample comprising 31 data points in the redshift range $z \in [0.01, 1.3]$ [104].
- *Observational Hubble data (OHD):* For OHD data set, we used the 43 measurements of the Hubble parameter values in the interval $0 < z < 2.5$, see [105].
- *Baryon acoustic oscillations and cosmic microwave background(BAO/CMB):* We use the combined BAO/CMB data from different observational missions [106]. We have taken 2 data points from SDSS(R) [107], 1 data point from the 6dF Galaxy survey [108] and BOSS CMASS [109] and 3 data points from the from WiggleZ survey [110]. We combine these results with the Planck 2015 [111].

In order to fit the free parameters of our model, we use the Markov chain Monte Carlo (MCMC) method by employing EMCEE python package [112]. In our analysis, we have considered the value of the Hubble constant as $H_0 = 67.8$ Km/sec/Mpc [111].

For our statistical analysis, we consider the maximum likelihood function $\mathcal{L}_{max} \propto e^{-\chi^2_{\rm tot}/2}$, where $\chi^2_{\rm tot} = \chi^2_{SNe} + \chi^2_{OHD} + \chi^2_{BAO/CMB}$. We calculate the best-fit values of the model parameters from the minimization of the $\chi^2_{\rm tot}$ function. The best fit values for this viscous HRDE model are $\xi_0 = 0.603^{+0.112}_{-0.164}$, $\xi_1 = -0.149^{+0.031}_{-0.047}$, $\alpha = 0.726^{+1.049}_{-0.389}$, $\omega_{de} = -0.136^{+0.092}_{-0.174}$, which are summarized in Table 1. In what follows, we use this best fit value to plot the relation between the parameters and discuss the GSL of thermodynamics for BH entropy, Barrow entropy and Casimir effect in this viscous model.

Using the best-fit values in Eq. (13), we plot a graph of q with redshift z as shown in Fig. 1. It shows that the viscous model has a transition from decelerated phase to accelerated phase. We obtain the present value of $q_0 = -0.386^{+0.361}_{-0.194}$ and the transition takes place at $z_{tr} = 0.753^{+0.78}_{-0.72}$, where z_{tr} is the transition redshift at which $q = 0$. In late time, it is observed that $q \to -1$. The evolution of Hubble parameter $H(z)$ with Hubble error bar of 43 data points is shown in Fig. 2. It can be observed the viscous HDRE model fits very well between Hubble error bar data through out the evolution and is in good agreement with the ΛCDM model.

In FRW universe, the bulk viscosity gives rise to the generation of the local entropy [113, 114]. The law of generation of the local entropy is given as

$$T\nabla_\nu s^\nu = \zeta(\nabla_\nu u^\nu)^2 = 9H^2\zeta, \tag{14}$$

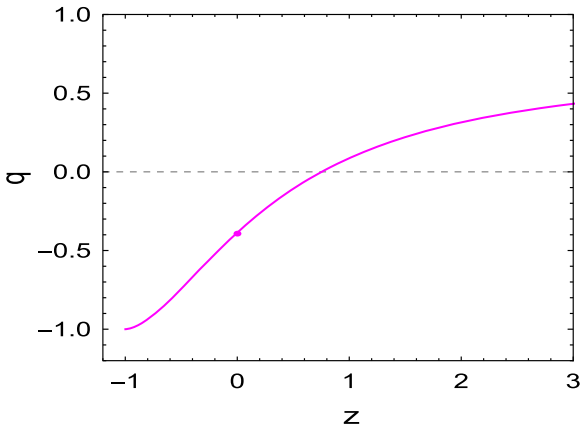


Fig. 1 The deceleration parameter-redshift relation using best-fit values obtained from $SNe + OHD + BAO/CMB$ data. A dot denotes the present value of q, i.e., q_0

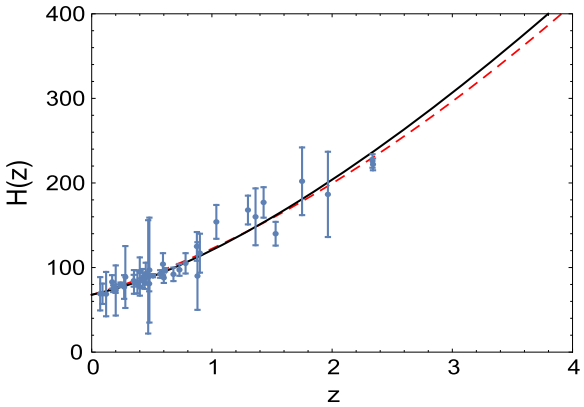


Fig. 2 The plot of $H(z)$ for Viscous HDRE model (dashed red curve) using the best-fit values. The observational 43 $H(z)$ data points are shown with error bars (grey colour). The variation of the Hubble function in the standard ΛCDM model is also represented as the Grey and black solid curves

where $\nabla_\nu s^\nu$ represents the rate of generation of entropy per unit volume and T is the temperature.

The second law of thermodynamics establishes the concept of entropy which predicts the direction of spontaneous processes, and determines whether they are irreversible or impossible, despite obeying the requirement of conservation of energy, which is established in the first law of thermodynamics. The change in entropy of a closed system is always positive. Thus, the second law of thermodynamics states that the change in entropy in an isolated system is greater than or equal to zero for a spontaneous process.

Using (14), the second law of thermodynamics (entropy law) can be expressed as

$$T\nabla_\nu s^\nu \geq 0, \tag{15}$$

which implies that

$$\zeta \geq 0. \tag{16}$$

To check the evolution of the total bulk viscosity, we obtain dimensionless bulk viscous parameter ξ as

$$\xi(a) = \xi_0 + \xi_1\left[\frac{\xi_0}{(1-\xi_1+2\alpha\omega_{de})} + \left(1 - \frac{\xi_0}{(1-\xi_1+2\alpha\omega_{de})}\right) a^{-\frac{3(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}}\right]. \tag{17}$$

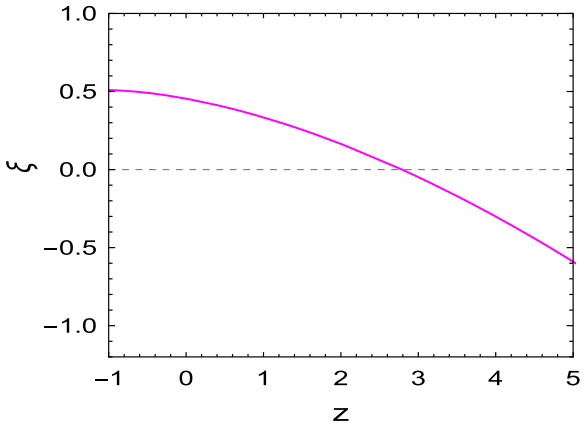


Fig. 3 The evolution of ξ with the respect to the scale factor a using best fit values of model parameters

Using best fit values, we plot the evolution of dimensionless bulk viscous parameter ξ with respect to cosmic redshift in Fig. 3. We observe that the trajectory of ξ evolves from negative to positive and hence the second law of thermodynamics (entropy law) violates in the early epoch and obeys in the later epoch. It can be found some recent arguments in literature which propose a solution to the cosmic initial entropy problem [115–118]

4 GSL of thermodynamics with Bekestein–Hawking entropy

In the previous section we have presented the cosmological scenario of the viscous HRDE model. The second law of thermodynamics suggests a link between entropy and the area of a black hole horizon. However, this law is not useful due to the disappearance of entropy near the exterior of black holes. Therefore, the generalized second law of thermodynamics is needed to present the second law of thermodynamics as valid. In the present section we proceed to study the thermodynamic aspects, and in particular the generalized second law (GSL) of thermodynamics [119,120] of viscous HRDE universe.

The GSL of thermodynamics states that the total entropy of the universe is always positive with time growth. The importance of this type of thermal study has been considered in Ref. [81]. It has been confirmed that gravity has a thermodynamical predisposition and the Einstein field equations of general relativity is just an equation of state for the spacetime. There is well-known conjecture that gravity and thermodynamics are related [121], and in particular one can show that the cosmological Friedmann equations can be expressed as the first law of thermodynamics, if we consider the universe as a thermodynamical system bounded by the apparent horizon [122] (for details, we refer to Refs. [123–128]). Similarly, performing the procedure in a reverse way, one can extract the Friedmann equations by applying the first law of thermodynamics. This procedure proves to be efficient both in general relativity as well as in a variety of modified gravity theories. Sheykhi [128] has constructed the cosmological field equations of the FRW universe with any special curvature, when the entropy associated with the apparent horizon is in the form of Barrow entropy.

To examine the thermodynamics of a cosmological models, one assumes the universe as a thermodynamical system with the apparent horizon surface being its boundary. The universe is assumed to be composed of HRDE and dark matter, and bounded by the horizon. GSL demands that the change in entropies of the components inside the universe when added together with that of horizon bounding the universe must be greater than or equal to zero. In other words, the total entropy of the interior of the universe plus the entropy of its horizon

should be a non-decreasing function of time [129]. This subject becomes more important when we discuss in the BH physics in which the usual entropy plus the BH horizon entropy is a non-decreasing function of time.

Concerning the universe horizon, that is the boundary of the thermodynamical system, we study the validity of GSL of thermodynamic of viscous HRDE in a universe enveloped by the apparent horizon. Sheykhi and Wang [120] have shown that for an accelerating universe the apparent horizon is a physical boundary from the thermodynamical point of view which is a marginally trapped surface with vanishing expansion given by [88,130]

$$\tilde{r}_A = \frac{1}{\sqrt{H^2 + \frac{k}{a^2}}} \tag{18}$$

where k quantifies the spatial curvature which is set to zero in the present work. Hence, in terms of the apparent horizon radius, the Friedmann equation (7) can be rewritten as

$$\frac{1}{\tilde{r}_A^2} = \frac{1}{3}(\rho_m + \rho_{de}) \tag{19}$$

In general, the apparent horizon $\tilde{r}_A$ is time-dependent. Hence, a change $d\tilde{r}_A$ in time interval dt will lead to a volume-change dV, while the energy and entropy of the universe fluids will change by dE and dS respectively. Let us first establish the mathematical formula for the entropy using the first law of thermodynamics [91,131]

$$TdS = dE + PdV, \tag{20}$$

where T, S, P, V and $E = \rho V$ represent the temperature, entropy, pressure, volume and internal energy of the considered system, respectively. Therefore, using (20) the dark matter and HRDE entropies read [90,94]

$$T_{de}dS_{de} = d(\rho_{de}V) + \tilde{p}_{de}dV = d(\rho_{de}V) + (p_{de} - 3\zeta H)dV, \tag{21}$$

$$T_m dS_m = d(\rho_m V) + p_m dV, \tag{22}$$

where T_{de} and T_m are the temperature of the viscous HRDE and dark matter, respectively, and $V = 4\pi\tilde{r}_A^3/3$ is the volume of the system bounded by the apparent horizon which gives $dV = 4\pi\tilde{r}_A^2 d\tilde{r}_A$.

Finally, the standard choice of the entropy associated to the apparent horizon is to use Bekestein–Hawking entropy [121,122,132]. It is defined as $S_h = 4\pi\tilde{r}_A^2/(4G)$, and since we use the unit $8\pi G = 1$ in this work, therefore, we obtain

$$S_h = 8\pi^2\tilde{r}_A^2. \tag{23}$$

According to the generalization of black-hole thermodynamics to a cosmological framework, the temperature of the horizon is related to its radius through [96]

$$T_h = \frac{1}{2\pi\tilde{r}_A}\left(1 - \frac{\dot{\tilde{r}}_A}{2H\tilde{r}_A}\right), \tag{24}$$

where $\frac{\dot{\tilde{r}}_A}{2H\tilde{r}_A} < 1$ ensures that the temperature is positive. The negative temperature is not physically acceptable and hence we define $T = |\frac{1}{\tilde{r}_A}\left(1 - \frac{\dot{\tilde{r}}_A}{2H\tilde{r}_A}\right)|/2\pi$.

Further, we assume that the thermal system bounded by the apparent horizon remains in equilibrium. After attaining equilibrium and the universe geometry becomes FRW, all the fluids in the universe acquire the same temperature [133], which is moreover equal

to the temperature of apparent horizon T_h otherwise the energy flow would deform the FRW geometry. This gives the temperature T_{de} and T_m of viscous HRDE and dark matter, respectively, inside the apparent horizon to be in equilibrium with the temperature T_h [92, 95,134], associated with the apparent horizon, i.e., $T_{de} = T_m = T_h = T$. Thus, we follow this assumption keeping in mind that our results hold at late times of the evolution of the universe

Now, Dividing (21) and (22) respectively by dt, we get

$$T\dot{S}_{de} = 4\pi\tilde{r}_A^2\left[(1+\omega_{de})\rho_{de} - 3\zeta H\right]\left(\dot{\tilde{r}}_A - \tilde{r}_A H\right), \tag{25}$$

$$T\dot{S}_m = 4\pi\tilde{r}_A^2\rho_m\left(\dot{\tilde{r}}_A - \tilde{r}_A H\right), \tag{26}$$

where

$$\dot{\tilde{r}}_A = \frac{1}{2}H\tilde{r}_A^3\left[(1+\omega_{de})\rho_{de} + \rho_m - 3\zeta H\right]. \tag{27}$$

Equation (27) can be obtained from Friedmann equation (7) by using (8) and (9). Equations (25) and (26) provide the entropy time-variation for the universe interior.

From (23), we obtain

$$\dot{S}_h = 16\pi^2\tilde{r}_A\dot{\tilde{r}}_A. \tag{28}$$

Using (24) and (27), Eq. (28) can be written as

$$T\dot{S}_h = 4\pi\tilde{r}_A^3 H\left[(1+\omega_{de})\rho_{de} + \rho_m - 3\zeta H\right]\left(1 - \frac{\dot{\tilde{r}}_A}{2\tilde{r}_A H}\right). \tag{29}$$

Now, we can calculate the total entropy time-variation. Adding Eqs. (25), (26) and (29), we find

$$T\left(\dot{S}_m + \dot{S}_{de} + \dot{S}_h\right) = \pi\tilde{r}_A^5 H\left[(1+\omega_{de})\rho_{de} + \rho_m - 3\zeta H\right]^2. \tag{30}$$

Equation (30) gives the expression for GSL of thermodynamics where viscous HRDE do not interact with dark matter. The right hand side of (30) can not be negative throughout the history of the universe which means that $\dot{S}_m + \dot{S}_{de} + \dot{S}_h \geq 0$ always hold.

Using (11), Eq. (30) can be written in terms of redshift z as

$$T(\dot{S}_m + \dot{S}_{de} + \dot{S}_h) = \frac{36\pi[-1+\xi_0+\xi_1-2\alpha\omega_{de}]^2}{(2+3\alpha\omega_{de})^2\left[1+\frac{\xi_0\{(1+z)^{\frac{-3(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}}-1\}}{(1-\xi_1+2\alpha\omega_{de})}\right]^2} \tag{31}$$

The RHS of (31) is always non-negative throughout the evolution of the universe, which means that $(\dot{S}_m + \dot{S}_{de} + \dot{S}_h) \geq 0$ always true (see, Fig. 4). This implies that the GSL of thermodynamics always hold for a spatially flat FRW universe filled with bulk viscous HRDE and pressureless dark matter in a region enveloped by the apparent horizon.

To check the thermodynamical equilibrium, we solve (31) to obtain $\ddot{S}_{\rm tot}$ as

$$\ddot{S}_{\rm tot} = -\Bigg[432\pi^2(-1+\xi_0+\xi_1-2\alpha\omega_{de})^2$$

$$\left\{6\xi_0^2 + \left(6+12\alpha\omega_{de}+4(2+3\alpha\omega_{de})\left[1+\frac{\xi_0\{(1+z)^{\frac{-3(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}}-1\}}{(1-\xi_1+2\alpha\omega_{de})}\right]^2\right.\right.$$

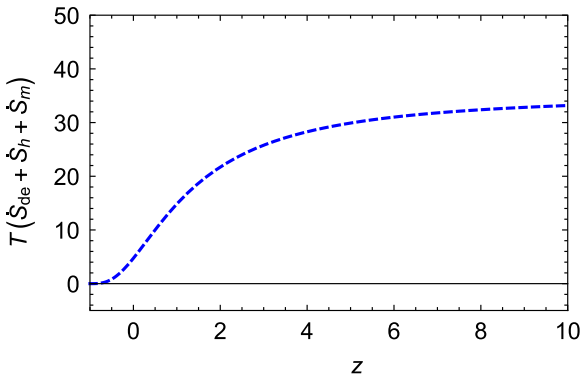


Fig. 4 Plot of $T(\dot{S}_m + \dot{S}_{de} + \dot{S}_h)$ with respect to z using best fit $\xi_0 = 0.603$, $\xi_1 = -0.149$, $\alpha = 0.726$ and $\omega_{de} = -0.136$ obtained from joint analysis of SNe, OHD and BAO/CMB data set

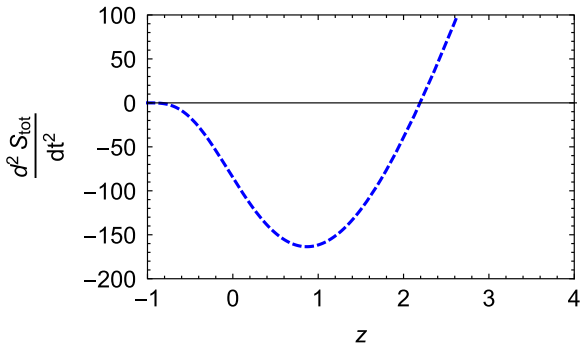


Fig. 5 Plot of $\ddot{S}_{\rm tot}$ with respect to z using best fit values of model parameters

$$
\begin{aligned}
&-3(5-\xi_1+8\alpha\omega_{de})\left[1+\frac{\xi_0\{(1+z)^{\frac{-3(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}}-1\}}{(1-\xi_1+2\alpha\omega_{de})}\right]-6\xi_1\Big)(1-\xi_1+2\alpha\omega_{de})\\
&+3\xi_0\Big(-4-8\alpha\omega_{de}\\
&+(5-\xi_1+8\alpha\omega_{de})\left[1+\frac{\xi_0\{(1+z)^{\frac{-3(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}}-1\}}{(1-\xi_1+2\alpha\omega_{de})}\right]+4\xi_1\Big)\Big\}\Big]\Big/\\
&\left[(2+3\alpha\omega_{de})^2\left[1+\frac{\xi_0\{(1+z)^{\frac{-3(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}}-1\}}{(1-\xi_1+2\alpha\omega_{de})}\right]^2\right.\\
&\left.\left[-3-6\alpha\omega_{de}+3\xi_0+(4+6\alpha\omega_{de})\left[1+\frac{\xi_0\{(1+z)^{\frac{-3(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}}-1\}}{(1-\xi_1+2\alpha\omega_{de})}\right]+3\xi_1\right]^2\right]
\end{aligned}
\tag{32}
$$

We plot the graph of $\ddot{S}_{\rm tot}$ for the observational values of model parameters in Fig. 5 which shows that $\ddot{S}_{\rm tot} > 0$ in early time whereas $\ddot{S}_{\rm tot} < 0$ in present time and in late time during the evolution. Hence, thermal equilibrium condition is satisfied for our model in present and late time at apparent horizon.

5 GSL of thermodynamics with Barrow entropy

In this section, we present the thermodynamics of GSL of viscous HRDE model using Barrow entropy. Recently, Barrow [89] modified the BH entropy caused by quantum-gravitational effects that deform the horizon, leading it to acquire a intricate, fractal structure on the black-hole surface and thus changing its actual horizon area. This in turn leads to a new black-hole entropy relation

$$S_B = \left(\frac{A}{A_0}\right)^{1+\Delta/2}, \tag{33}$$

where A is the usual horizon area and A_0 is the Planck area. The quantum gravitational perturbation is represented by the new exponent Δ and it is bounded as $0 \leq \Delta \leq 1$. The value $\Delta = 0$ corresponds to the standard BH entropy which has been discussed in previous section where as $\Delta = 1$ gives the maximal deformation. Some efforts have been made to disclose the influence of Barrow entropy in the cosmological setups (see, e.g., [128] and references therein).

The Barrow black-hole entropy associated to the apparent horizon with the standard area $A = 4\pi\tilde{r}_A^2$ is given by [135]

$$S_B = \gamma \tilde{r}_A^{\Delta+2}, \tag{34}$$

where $\gamma \equiv (4\pi/A_0)^{1+\Delta/2}$. Differentiating (34), we obtain

$$T\dot{S}_B = \frac{1}{4\pi}\gamma(2+\Delta)\tilde{r}_A^{3+\Delta} H\left[(1+\omega_{de})\rho_{de} + \rho_m - 3\zeta H\right]\left(1 - \frac{\dot{\tilde{r}}_A}{2\tilde{r}_A H}\right). \tag{35}$$

Now, adding (25), (26) and (35), we obtain

$$\begin{aligned} T\left(\dot{S}_m + \dot{S}_{de} + \dot{S}_B\right) = &\left(4\pi - \frac{\gamma(2+\Delta)}{12\pi}\tilde{r}_A^{\Delta}\right) H\tilde{r}_A^5\left[(1+\omega_{de})\rho_{de} + \rho_m - 3\zeta H\right]^2 \\ &+ \left(\frac{\gamma(2+\Delta)}{4\pi}\tilde{r}_A^{\Delta} - 4\pi\right) H\tilde{r}_A^3\left[(1+\omega_{de})\rho_{de} + \rho_m - 3\zeta H\right]. \end{aligned} \tag{36}$$

In terms of redshift z, Eq. (36) can be written as

$$\begin{aligned} T\left(\dot{S}_m + \dot{S}_{de} + \dot{S}_B\right) = -\Bigg[&3(1-\xi_0-\xi_1+2\alpha\omega_{de}) \\ &\left\{32\pi^2\left[(2+3\alpha\omega_{de})\left[1 + \frac{\xi_0\{(1+z)^{\frac{-3(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}} - 1\}}{(1-\xi_1+2\alpha\omega_{de})}\right]\right.\right. \\ &\left.\left. + 3(-1+\xi_0+\xi_1-2\alpha\omega_{de})\right]\right. \\ &- \delta(2+\Delta)\left\{\frac{(1+z)^{\frac{-3(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}}}{\left[1 + \frac{\xi_0\{(1+z)^{\frac{-3(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}} - 1\}}{(1-\xi_1+2\alpha\omega_{de})}\right]}\right\}^{\Delta} \\ &\left[(4+6\alpha\omega_{de})\left[1 + \frac{\xi_0\{(1+z)^{\frac{-3(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}} - 1\}}{(1-\xi_1+2\alpha\omega_{de})}\right]\right. \\ &\left.\left.\left. + 3(-1+\xi_0+\xi_1-2\alpha\omega_{de})\right]\right\}\Bigg]\Bigg/ \end{aligned}$$

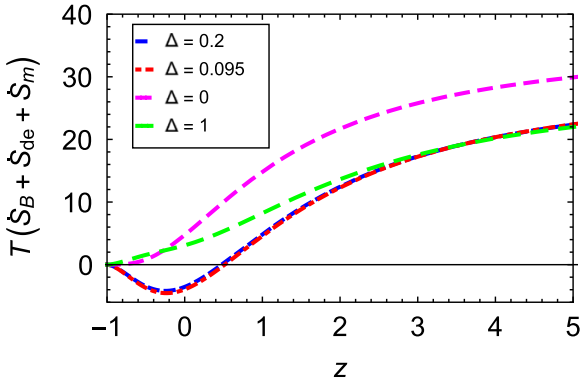


Fig. 6 Plot of $T(\dot{S}_m+\dot{S}_{de}+\dot{S}_B)$ with respect to z for different Δ values using best fit $\xi_0 = 0.603, \xi_1 = -0.149$, $\alpha = 0.726$ and $\omega_{de} = -0.136$ obtained from joint analysis of SNe, OHD and BAO/CMB data set along with some fixed values $A_0 = 1$ and $\gamma = 3.54$

$$\left[4\pi(2+3\alpha\omega_{de})^2\left[1+\frac{\xi_0\{(1+z)^{\frac{-3(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}}-1\}}{(1-\xi_1+2\alpha\omega_{de})}\right]^2\right] \tag{37}$$

In the case of standard Bekestein–Hawking black-hole thermodynamics, i.e., where $\Delta = 0$, we recover the results as discussed in previous section. Thus, the GSL of thermodynamic is always valid in a universe filled with matter and viscous HRDE and governed by general relativity.

However, in case where Barrow entropy $\Delta \neq 0$ is used, the total entropy is not necessarily a non-decreasing function of time, and hence GSL can be violated. The reason is that the horizon contributes with a changed $\dot{S}_B$, which does not satisfy certain conditions. Therefore, Barrow entropy leads to the conditional violation of GSL of thermodynamics, depending on the evolution of the universe. In Fig. 6, we plot the trajectory to discuss the behavior of Eq. (37). It can be observed that the total change in entropy violates the condition of GSL of thermodynamics for $\Delta = 0.2$ [136] and $\Delta = 0.095$ [137] in late time whereas holds in early time. However, GSL is always valid for $\Delta = 0$ and $\Delta = 1$. Therefore, the violation of GSL depends on the value of Δ in Barrow entropy.

6 Viscous HRDE with the Casimir effect

In this section, we present the GSL of thermodynamics of viscous HRDE due to the Casimir effect [96]. We consider that the Casimir energy E_c is inversely proportional to the apparent horizon radius, that is,

$$E_c = \frac{c}{\tilde{r}_A}, \tag{38}$$

where c is a constant which is assumed to be very small as compared to the unity. The basic property is that the Casimir energy is a measure of the stress in the region interior to the "shell" as compared with the unstressed region on the outside. It can be observed from (38) that the effect of the Casimir energy is largest in early-time when $\tilde{r}_A$ is small. However, it fades away when radius tends to infinity.

The corresponding Casimir pressure is

$$p_c = \frac{-1}{4\pi\tilde{r}_A^2}\frac{\partial E_c}{\partial \tilde{r}_A} = \frac{c}{4\pi\tilde{r}_A^4} \tag{39}$$

Thus, the Casimir energy density evolves as $\rho_c \propto \tilde{r}_A^{-4}$. Now, the energy conservation equation for the Casimir energy takes the form

$$\dot{\rho}_c + 3H(1+\omega_c)\rho_c = 0 \tag{40}$$

where $\omega_c = p_c/\rho_c$ is the EoS parameter of the Casimir energy. As we know the relation $\rho_c = E_c/V = 3c/4\pi\tilde{r}_A^4$, where $V = 4\pi\tilde{r}_A^3/3$, we have $\omega_c = 1/3$.

The Friedmann equation (7) can be rewritten as

$$\frac{1}{\tilde{r}_A^2} = \frac{1}{3}(\rho_m + \rho_{de} + \rho_c) \tag{41}$$

Differentiating (41) with respect to the cosmic time and using (8), (9) and (40), we get

$$\dot{\tilde{r}}_A = \frac{1}{2}\tilde{r}_A^3 H\left[(1+\omega_{de})\rho_{de} + \frac{4}{3}\rho_c + \rho_m - 3\zeta H\right] \tag{42}$$

Now, using (24) and (42), the change in entropy of apparent horizon, given in (29), can take the form

$$T\dot{S}_h = 4\pi\tilde{r}_A^3 H\left[(1+\omega_{de})\rho_{de} + \frac{4}{3}\rho_c + \rho_m - 3\zeta H\right]\left(1 - \frac{\dot{\tilde{r}}_A}{2\tilde{r}_A H}\right) \tag{43}$$

The Gibbs equation for effect of the Casimir energy is given by

$$T dS_c = d(\rho_c V) + p_c dV, \tag{44}$$

which gives

$$T\dot{S}_c = \frac{16}{3}\pi\tilde{r}_A^2\rho_c(\dot{\tilde{r}}_A - \tilde{r}_A H) \tag{45}$$

It is to be noted that the temperatures of the apparent horizon and interior of the universe are assumed to be equal in FRW model. Now, we are in position to find the total entropy which can be obtained by adding (25), (26), (43) and (45), that is,

$$T\dot{S}_{\text{tot}} = 2\pi\tilde{r}_A^2\left[(1+\omega_{de})\rho_{de} + \rho_m + \frac{4}{3}\rho_c - 3\zeta H\right]\dot{\tilde{r}}_A, \tag{46}$$

where $\dot{S}_{\text{tot}} = \dot{S}_{de} + \dot{S}_m + \dot{S}_c + \dot{S}_h$. Using (42), Eq. (46) finally gives

$$T\dot{S}_{\text{tot}} = \pi\tilde{r}_A^5 H\left[(1+\omega_{de})\rho_{de} + \rho_m + \frac{4}{3}\rho_c - 3\zeta H\right]^2, \tag{47}$$

which gives $\dot{S}_{\text{tot}} \geq 0$. Therefore, the GSL of thermodynamics including the Casimir effect always holds with viscous HRDE and dark matter in a region enclosed by the apparent horizon. The above equation in terms of z can be written as

$$\begin{aligned} T\dot{S}_{\text{tot}} = \Bigg[& (1+z)^{\frac{12(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}} \\ & \Bigg\{ cH_0^2\xi_0^2(2+3\alpha\omega_{de})\Big(1-\xi_1+2\alpha\omega_{de}+\xi_0\{(1+z)^{\frac{-3(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}}-1\}\Big)^3 \\ & - 6\pi\xi_0^2(1-\xi_1+2\alpha\omega_{de})^3(-1+\xi_0+\xi_1-2\alpha\omega_{de})(1+z)^{\frac{-6(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}} \Bigg\}^2 \Bigg] \Big/ \end{aligned}$$

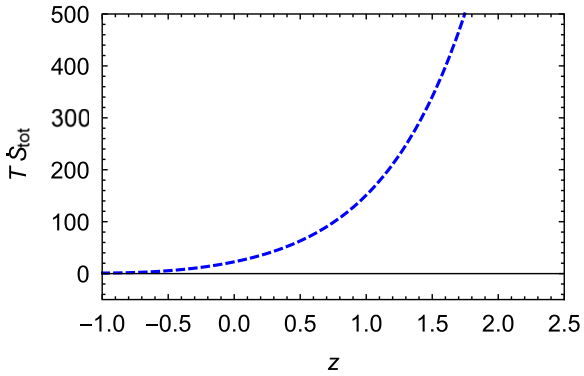


Fig. 7 Plot of $\dot{S}_{\rm tot}$ with respect to z using best fit $\xi_0 = 0.603$, $\xi_1 = -0.149$, $\alpha = 0.726$ and $\omega_{de} = -0.136$ obtained from joint analysis of SNe, OHD and BAO/CMB data set

$$\Bigg[\pi\xi_0^4(2+3\alpha\omega_{de})^2(1-\xi_1+2\alpha\omega_{de})^6\Bigg[1+\frac{\xi_0}{(1-\xi_1+2\alpha\omega_{de})} \{(1+z)^{\frac{-3(1-\xi_1+2\alpha\omega_{de})}{(2+3\alpha\omega_{de})}}-1\}\Bigg]^2\Bigg], \tag{48}$$

From Eq. (48), we observe that the RHS can never be negative throughout the evolution of the universe (See, Fig. 7). Hence, the GSL of thermodynamics is always valid for our model taking into account the Casimir effect.

7 Conclusions

In this paper, we have discussed the validity of the GSL of thermodynamics for a flat FRW universe filled with pressureless dark matter and viscous HRDE. In a homogeneous and isotropic universe bulk viscosity is the unique dissipative effect which is capable to modify the background dynamics. The dissipative effect caused by the bulk viscosity can be redefined by the effective pressure. As the explicit form of the bulk viscosity is still unknown, we have assumed the bulk viscous coefficient as $\zeta = \zeta_0 + \zeta_1 H$, to find the analytical expressions for the evolution of Hubble parameter, the scale factor and the bulk viscous coefficient. The observational test has been carried out by using joint analysis of observational data: SNe, OHD and BAO/CMB, to obtained the best fit values of model parameters. It can be observed that the viscous HRDE model undergoes early deceleration and then late-time acceleration. In the absence of bulk viscosity, the HRDE model always decelerates or accelerates. Thus, we can find a transition from decelerated phase to accelerated phase in the presence of bulk viscosity. In what follows, we summarize the main results.

To discuss the evolution of the viscous HRDE model, we have plotted $q - z$ trajectory and $H(z)$ with error bars as shown in Figs. 1 and 2, respectively. It has been observed that the model shows transition from decelerated phase to accelerated phase at transition redshift $z_{tr} = 0.753^{+0.78}_{-0.72}$. The present value of q is $q_0 = -0.386^{+0.361}_{-0.194}$. The viscous HRDE model is in good agreement with ΛCDM model and observationally fit the $H(z)$ error bars data points very well. In Fig. 3 we plot the viscous coefficient with the redshift using the best fit values. According to GSL, it requires that $\zeta > 0$, which assures the positive entropy production. From Fig. 3, it has been observed that it violates in early phase of evolution whereas it holds in late time of evolution.

Further, we have discussed the GSL of thermodynamics with Bekestein–Hawking entropy and have examined the total change in entropy. In Fig. 4, we have plotted the evolution of the total change in entropy, including the entropy associated with the apparent horizon and the entropy of the viscous HRDE and dark matter inside the apparent horizon. It has been observed that the total change in entropy is always a non-decreasing function of time and thus the GSL of thermodynamics is always hold with the best fit values of free parameters of a viscous HRDE model enclosed by apparent horizon. The condition of thermal equilibrium is satisfied for our model in late time whereas fails to hold in early time in the universe enclosed by the apparent horizon as shown in Fig. 5.

We extend our investigated to examine the validity of the GSL of thermodynamics with Barrow entropy for the horizon entropy which arises from the fact that the black-hole surface may be deformed due to quantum-gravitational effects, quantified by a new exponent Δ. We have calculated the time evolution of total entropy in a universe filled with dark matter and viscous HRDE fluids, as well as the corresponding quantity for the apparent horizon. It has been found that in case of $\Delta = 0$, we recover the results of Bekestein–Hawking entropy and the GSL is valid. However, in case of $\Delta \neq 0$, this may not be true. This means that the total entropy may not be always a non-decreasing function of time and hence the GSL may be violated depending on the exponent Δ. In Fig. 6, we have observed that for $\Delta = 0$ and $\Delta = 1$, the total entropy is always non-decreasing function of time where as for some values of Δ, like $\Delta = 0.2$ and $\Delta = 0.095$, the total entropy is positive in early time whereas negative in late time of the evolution and hence the GSL is valid in early time and not valid in late time of the evolution. Hence, the validity of the GSL of thermodynamics with Barrow entropy depends on the exponent Δ.

Finally, we have examined the GSL of thermodynamics of viscous HRDE model by taking into account of the Casimir effect. It has been found that the change in total entropy is always positive throughout the evolution of the universe as shown in Fig. 7. Therefore, the GSL of thermodynamics due to Casimir effect is always valid with the best fit values of free parameters in viscous HRDE model.

Acknowledgements The authors are highly appreciate the anonymous referee for very constructive comments, which helped us to improve the manuscript significantly. One of the authors, AK is grateful to University Grant Commission, New Delhi for providing SRF under CSIR-UGC NET-JRF scheme.

Appendix A: data sets

1. Supernovae type Ia (SNe Ia)

For the supernova type Ia (SNe Ia), we use 31 binned data points within the range of redshift $z \propto [0.01, 1.3]$. The χ^2 for SNe Ia is defined as

$$\chi^2_{SNe} = r^t C_b^{-1} r \tag{49}$$

in which

$$r = \mu_b - M - 5 \log_{10} d_L \tag{50}$$

here μ_b is the observational distance modulus for a supernova at redshift z, M is a free normalization parameter and C_b is the covariance matrix of μ_b, see Table F.2 in [104]. In the

above equation d_L is the dimensionless luminosity distance which is defined by

$$d_L = \frac{c(1+z)}{H_0}\int_0^z \frac{\mathrm{dz}'}{H(z',\theta)} \tag{51}$$

where θ represents the set of model parameters and c is the speed of light.

2. Observational hubble data (OHD)

For the study of the cosmic expansion history, the measurement of Hubble parameter $H(z)$ got much attention from the researchers due to its model independent nature. In this paper, we use the 43 measurements of the Hubble parameter values in the redshift range $0 < z < 2.5$ [105]. The χ^2 for OHD is given by

$$\chi^2_{OHD} = \sum_{i=1}^{n} \frac{[H(z_i) - H_{obs}(z_i,\theta)]^2}{\sigma_i^2}, \tag{52}$$

where $H(z_i)$ and $H_{obs}(z_i)$ represents theoretical and observed values respectively and σ_i^2 the standard deviation of each $H_{obs}(z_i)$.

3. Baryon acoustic oscillations and cosmic microwave background (BAO/CMB)

In the context of BAO/CMB, we can write the χ^2 function as [106]

$$\chi^2_{\mathrm{BAO/CMB}} = A^T C^{-1} A \tag{53}$$

where C^{-1} is the inverse of the covariance matrix [106] and A is the matrix

$$A = \begin{bmatrix} \frac{d_A(z_*,\theta)}{D_v(0.106,\theta)} - 30.84 \\ \frac{d_A(z_*,\theta)}{D_v(0.35,\theta)} - 10.33 \\ \frac{d_A(z_*,\theta)}{D_v(0.57,\theta)} - 6.72 \\ \frac{d_A(z_*,\theta)}{D_v(0.44,\theta)} - 8.41 \\ \frac{d_A(z_*,\theta)}{D_v(0.6,\theta)} - 6.66 \\ \frac{d_A(z_*,\theta)}{D_v(0.73,\theta)} - 5.43 \end{bmatrix}$$

here, $D_v(z,\theta)$ 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 comoving angular diameter, $\mathrm{d}_A(z,\theta)$ is defined as

$$d_A(z_*,\theta) = c\int_0^{z_*} \frac{dz'}{H(z',\theta)} \tag{54}$$

where z_* indicates the photons decoupling redshift and hold the value $z_* = 1090$ as per the Planck 2015 results [111]. We have taken the correlation coefficient from Ref. [138].

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. D.N. Spergel et al., Astrophys. J. Suppl. Ser. **148**, 175 (2003)
5. D.N. Spergel et al., Astrophys. J. Suppl. Ser. **170**, 377 (2007)

6. M. Tegmark et al., Phys. Rev. D **69**, 103501 (2004)
7. J. Silk, Ann. Phys. (Berlin) **15**, 75 (2006)
8. L. Bergstrom, Ann. Phys. (Berlin) **524**, 479 (2012)
9. C.S. Frenk, S.D.M. White, Ann. Phys. (Berlin) **524**, 507 (2012)
10. W. Weinberg, Rev. Mod. Phys. **61**, 1 (1989)
11. P.J.E. Peebles, B. Ratra, Astrophys. J. Lett. **325**, L17 (1988)
12. R.R. Caldwell, R. Dave, P.J. Steinhardt, Phys. Rev. Lett. **80**, 1582 (1998)
13. B. Ratra, P.J.E. Peebles, Phys. Rev. D **37**, 3406 (1988)
14. C. Armendariz-Picon, V.F. Mukhanov, P.J. Steinhardt, Phys. Rev. Lett. **85**, 4438 (2000)
15. T. Chiba, T. Okabe, M. Yamaguchi, Phys. Rev. D **62**, 023511 (2000)
16. A. Sen, J. High Energy Phys. **07**, 065 (2002)
17. T. Padmanabhan, Phys. Rev. D **66**, 021301 (2002)
18. R.R. Caldwell, Phys. Lett. B **545**, 23 (2002)
19. A. Kamenshchik, U. Moschella, V. Pasquier, Phys. Lett. B **511**, 265 (2001)
20. S.D.H. Hsu, Phys. Lett. B **594**, 13 (2004)
21. M. Li, Phys. Lett. B **603**, 1 (2004)
22. R.G. Cai, Phys. Lett. B **657**, 228 (2007)
23. H. Wei, R.G. Cai, Phys. Lett. B **660**, 113 (2008)
24. G. 't Hooft, arXiv:gr-qc/9310026 (1993)
25. L. Susskind, J. Math. Phys. **36**, 6377 (1995)
26. R. Bousso, Class. quantum Grav. **17**, 997 (2000)
27. F. Karolyhazy, Nuovo Cimento **42**, 390 (1966)
28. A.G. Cohen, D.B. Kaplan, A.E. Nelson, Phys. Rev. Lett. **82**, 4971 (1999)
29. C. Gao, F. Wu, X. Chen, Phys. Rev. D **79**, 043511 (2009)
30. C.J. Feng, Phys. Lett. B **670**, 231 (2008)
31. L. Xu, W. Li, J. Lu, Mod. Phys. Lett. A **24**, 1355 (2009)
32. C.J. Feng, X.Z. Li, Phys. Lett. B **679**, 151 (2009)
33. C.J. Feng, Phys. Lett. B **676**, 168 (2009)
34. C.J. Feng, X.Z. Li, Phys. Lett. B **680**, 184 (2009)
35. C.J. Feng, X. Zhang, Phys. Lett. B **680**, 399 (2009)
36. C.J. Feng, Phys. Lett. B **672**, 94 (2009)
37. C.J. Feng, X.Z. Li, Phys. Lett. B **680**, 355 (2009)
38. X. Zhang, Phys. Rev. D **79**, 103509 (2009)
39. S. Chattopadhyay, Eur. Phys. J. Plus **126**, 130 (2011)
40. S. Chattopadhyay, Proc. Natl. Acad. Sci., India, Sect. A Phys. Sci. **84**, (2014) 87
41. K.Y. Kim, H.W. Lee, Y.S. Myung, Gen. Relativ. Gravit. **43**, 1095 (2011)
42. I. Durán, D. Pavón, Phys. Rev. D **83**, 023504 (2011)
43. H. Saadat, Int. J. Theor. Phys. **51**, 731 (2012)
44. T.F. Fu et al., Eur. Phys. J. C **72**, 1932 (2012)
45. L.P. Chimento, M.G. Richarte, Eur. Phys. J. C **73**, 2352 (2013)
46. S. Del Campo et al., Phys. Rev. D **87**, 123002 (2013)
47. C. Eckart, Phys. Rev. **58**, 919 (1940)
48. L.D. Landau, E.M. Lifshitz, *Fluid Mechanics*, vol. 6 (Butterworth Heinermann Ltd., Oxford, 1987)
49. W. Israel, J.M. Stewart, Phys. Lett. A **58**, 213 (1976)
50. I. Brevik, O. Gorbunova, Gen. Relativ. Gravit. **37**, 2039 (2005)
51. W. Zimdahl, D.J. Schwarz, A.B. Balakin, D. Pavon, Phys. Rev. D **64**, 063501 (2001)
52. J.R. Wilson, G.J. Mathews, G.M. Fuller, Phys. Rev. D **75**, 043521 (2007)
53. Ø. Grøn, Astrophys. Space Sci. **173**, 191 (1990)
54. C.P. Singh, Pramana. J. Phys. **71**, 33 (2008)
55. I. Brevik, S.D. Odintsov, Phys. Rev. D **65**, 067302 (2002)
56. C.P. Singh, S. Kumar, A. Pradhan, Class. Quantum Grav. **24**, 455 (2007)
57. J.C. Fabris, S.V.B. Goncalves and R.de S. Ribeiro, Gen. Relativ. Grav. **38**, (2006) 495
58. M.G. Hu and X.H. Meng, Phys. Lett.B **635**, (2006) 186
59. B. Li, J.D. Barrow, Phys. Rev. D **79**, 103521 (2009)
60. A. Avelino, U. Nucamendi, JCAP **04**, 006 (2009)
61. A. Avelino, U. Nucamendi, JCAP **08**, 009 (2010a)
62. A. Avelino, U. Nucamendi, JCAP **09**, 1008 (2010b)
63. X.H. Meng, X. Dou, Commun. Theor. Phys. **52**, 377 (2009)
64. C.P. Singh, P. Kumar, Eur. Phys. J. C **74**, 3070 (2014)
65. A. Sasidharan, T.K. Mathew, Eur. Phys. J. C **75**, 348 (2015)

66. D. Wang, Y.-J. Yan, X.-H. Meng, Eur. Phys. J. C **77**, 660 (2017)
67. C.P. Singh, A. Kumar, Eur. Phys. J. Plus **133**, 312 (2018)
68. C.P. Singh, A. Kumar, Mod. Phys. Lett. A **33**(38), 1850225 (2018)
69. C.P. Singh, M. Srivastava, Eur. Phys. J. C **78**, 190 (2018)
70. C.P. Singh, A. Kumar, Astrophys. Space Sci. **364**, 94 (2019)
71. C.P. Singh, S. Kaur, Astrophys. Space Sci **365**, 2 (2020)
72. L. Herrera-Zamorano, A. Hernandez-Almada and Miguel A. Garcia-Aspeitia, Eur. Phys. J. C **80**, (2020) 637
73. R. Cardenas, T. Gonzales, Y. Leira, O. Martin, I. Quiros, Phys. Rev. D **67**, 083501 (2003)
74. R.A. El-Nabulsi, Phys. Lett. B **619**, 26 (2005)
75. R.A. El-Nabulsi, Chin. Phys. Lett. **23**, 1124 (2006)
76. N. Ahmadi, M. Nouti-Zonoz, Class. Quantum Grav. **25**, 135008 (2008)
77. T. Biswas, A. Mazumdar, Phys. Rev. D **80**, 023519 (2009)
78. R.A. El-Nabulsi, Comm. Theor. Phys. **54**, 16 (2010)
79. L. Visinelli, S. Vagnozzi, U. Danielsson, Symmetry **11**(8), 1035 (2019)
80. R. Calderon, R. Gannouji, B. L'Huillier, D. Polaski, Phys. Rev. D **103**, 023526 (2021)
81. J.D. Bekenstein, Phys. Rev. D **7**, 2333 (1973)
82. J.D. Bekenstein, Phys. Rev. D **9**, 3292 (1974)
83. J.D. Bekenstein, Phys. Rev. D **12**, 3077 (1975)
84. S.W. Hawking, Commun. Math. Phys. **43**, 199 (1975)
85. G.W. Gibbons, S.W. Hawking, Phys. Rev. D **15**, 2738 (1977)
86. R.G. Cai, L.M. Cao, Y.P. Hu, Class. Quant. Grav. **26**, 155018 (2009)
87. J. Sola Peracaula and H. Yu, Gen. Relativ. Grav. **52**, (2020) 17
88. D. Bak, S.J. Rey, Class. Quantum Grav. **17**, L83 (2000)
89. J.D. Barrow, Phys. Lett. B **808**, 135643 (2020)
90. B. Wang, Y. Gong, E. Abdalla, Phys. Rev. D **74**, 083520 (2006)
91. G. Izquierdo, D. Pavon, Phys. Lett. B **633**, 420 (2006)
92. M. Akbar, R.G. Cai, Phys. Rev. D **75**, 084003 (2007)
93. J. Zhou, B. Wang, Y. Gong, E. Abdalla, Phys. Lett. B **652**, 86 (2007)
94. Y. Gong, B. Wang, A. Wang, Phys. Rev. D **75**, 123516 (2007)
95. A. Sheykhi, Class. Quantum Grav. **27**, 025007 (2010)
96. M.R. Setare, A. Sheykhi, Int. J. Mod. Phys. D **19**, 1205 (2010)
97. J. Dutta, S. Chakraborty, Mod. Phys. Lett. A **25**, 3069 (2010)
98. K. Karami, S. Ghaffari, M.M. Soltanzadeh, Class. Quant. Grav. **27**, 205021 (2010)
99. K. Karami, A. Sheykhi, N. Sahraei, Eur. Phys. Lett. **93**, 29002 (2011)
100. M. Jamil, E.N. Saridakis, M.R. Setare, JCAP **11**, 032 (2011)
101. U. Debnath, Astrophys. space Sci. **333**, 503 (2012)
102. P. Praseetha, T.K. Mathew, Class. Quant. Grav. **31**, 185012 (2014)
103. P. Lore-Anguilar, E. Garci-Berro, J. Isern, Y.A. Kubyshin, Class. Qunatum Grav. **20**, 3885 (2003)
104. M. Betoule et al., Astron. Astrophys. **568**, A22 (2014)
105. S.L. Cao et al., Eur. Phys. J. C **78**, 170 (2018)
106. M.V. dos Santos, R.R.R. Reis, I. Waga, J. Cosm. Astrop. Phys. **02**, 066 (2016)
107. N. Padmanabhan et al., Mon. Not. R. Astron. Soc. **427**, 2132 (2012)
108. F. Beutler et al., Mon. Not. R. Astron. Soc. **416**, 3017 (2011)
109. L. Anderson et al., Mon. Not. R. Astron. Soc. **441**, 24 (2014)
110. C. Blake et al., Mon. Not. R. Astron. Soc. **425**, 405 (2012)
111. P.A.R. Ade et al., Astron. Astrophys. **594**, 13 (2016)
112. D. Foreman-Mackey et al., PASP **125**, 306 (2013)
113. S. Weinberg, Astrophys. J. **168**, 175 (1971)
114. S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (John Wiley and Sons Inc., New York, USA, 1972)
115. R.A. El-Nabulsi, Proc. Natl. Acad. Sci.: India Sect. A Phys. Sci. **85**, 395 (2015)
116. M.N. Bera, A. Riera, M. Lewenstein, A. Winter, Nat. Comm. **8**, 2180 (2017)
117. R.A. El-Nabulsi, Continuum Mech. Thermodyn. **30**, 889 (2018)
118. A.N. Tawfik, C. Greiner, Entropy **23**(3), 295 (2021)
119. M.R. Setare and E.C. Vagnas, Phys. Lett.B **666**, (2007) 111
120. A. Sheykhi, B. Wang, Phys. Lett. B **678**, 434 (2009)
121. T. Padmanabhan, Phys. Rept. **406**, 49 (2005)
122. R.G. Cai, S.P. Kim, J. High Energy Phys. **0502**, 050 (2005)
123. R.A. El-Nabulsi, Chin. Phys. Lett. **26**, 090401 (2009)

124. R.A. El-Nabulsi, Astrophys. space Sci. **325**, 277 (2010)
125. U. Debnath, M. Jamil, S. Chattopadhyay, Int. J. Theor. Phys. **51**, 812 (2012)
126. Y. Kang, S.-W. Kim, J. Korean Phys. Soc. **73**, 1 (2018)
127. R.A. El-Nabulsi, J. Korean Phys. Soc. **78**, 1800 (2021)
128. A. Sheykhi, Phys. Rev. D **103**, 123503 (2021)
129. J.D. Barrow, Nucl. Phys. B **310**, 743 (1988)
130. S.A. Hayward, S. Mukohyama, M. Ashworth, Phys. Lett. A **256**, 347 (1999)
131. E.K. Li, Gen. Relativ. Grav. **47**, 136 (2015)
132. T. Jacobson, Phys. Rev. Lett. **75**, 1260 (1995)
133. M.R. Setare, J. Cosmol. Astropart. Phys. **07**(01), 023 (2007)
134. G. Izquierdo, D. Pavón, Phys. Lett. B **639**, 1 (2006)
135. E.N. Saridakis and S. Basilakos, arXiv:2005.08258 [gr-qc]
136. E.N. Saridakis, arXiv:2006.01105 [gr-qc]
137. F.K. Anagnostropoulos, S. Basilakos, E.N. Saridakis, arXiv:2005.10302 [gr-qc]
138. G. Hinshaw et al., Astrophys. J. Suppl. Ser. **208**, 19 (2013)

The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/0368-492X.htm

The link between employee career orientation, career management practices, citizenship behavior and gender: a moderated mediation model

Received 31 May 2021
Revised 23 June 2021
20 July 2021
Accepted 8 August 2021

Meha Joshi, Girish Chandra Maheshwari and Rajan Yadav
Delhi School of Management, Delhi Technological University, Delhi, India

Abstract

Purpose – This study aims to add to the current understanding of mediation and moderation processes through which employee career orientation (CO) is linked with organizational citizenship behavior (OCB).

Design/methodology/approach – Data were collected from 214 employees working in Delhi and NCR of India. Structural equation modeling (SEM) was used for testing moderated mediation and establishing linkages between CO, CMP and OCB. Drawing on the social exchange theory, our model posits that the effect of CO on the outcome variable OCB is mediated by career management practices (CMPs) and the CMP-mediated relationship between the two is moderated by the gender of employees.

Findings – Overall, data from 214 employees from service organizations in India support the model. This suggests that the enactment of OCB as a consequence of CO and OCB is largely dependent on the gender of employees. The relationship was observed in such a way that for women employees, CMP will have a stronger influence on the CO-OCB relationship. According to bootstrap results, upon the addition of CMP as a mediator, the main effect of CO on OCB among male employees was significant but dropped from the Beta value of 0.281 to 0.196. However, adding CMP as a mediator among women employees caused the CO-OCB relationship to become insignificant (Beta = 0.124; LLCI = 0.415; ULCI = −0.127; p = 0.420), highlighting that CMP would have a more substantial influence on the CO-OCB relationship.

Originality/value – This study explains the mediational role of CMP in the relationship between CO and OCB (explaining how the employees with new CO can trigger the role of CMP, and consequently, CMP can help them enact OCB) and how the gender of employees moderate the mediated impact of CMP in the relationship between CO and OCB (explaining how the mediated relationship varies across genders). The novelty of the study lies in exploring such a relationship that has not been studied so far.

Keywords Human resource management, Behavior, Management, Modeling, Measurement, Psychology

Paper type Research paper

1. Introduction

Rapid changes in society and technological advancements have posed several challenges to organizations for motivating and retaining people at work by judiciously using career management systems on the one hand. At the same time, on the other hand, these influences point toward a growing awareness of an individual's responsibility for career development (Ronn, 2010). In the last two decades, the demise of traditional careers has been witnessed (Rodrigues *et al.*, 2021), and the world saw the dawn of a "new era." The new era is characterized by a new career orientation (CO) known as "protean," beholding self-directedness and value-drive in one's career. Due to it, the change in the nature of the psychological contract between the employers and employees becomes evident.

The changes in relationships have compelled organizations to rethink employees' career management as they significantly affect some crucial outcomes at work, such as employees' commitment toward the organization (De Vos and Cambré, 2017). Thus, in this era, OCB can be a stimulant for positive behaviors at work and realize organizational goals (Rafie, 2017). Also, it is conclusively proved that OCB can have tremendous effects on productivity and

Kybernetes
© Emerald Publishing Limited
0368-492X
DOI 10.1108/K-05-2021-0429

work outcomes. Moreover, OCB has a negative association with adverse consequences at work, such as stress (Tziner and Sharoni, 2014), counterproductive work behavior (Zhang *et al.*, 2019) and cynicism (Randy Evans *et al.*, 2010). Career theorists argue that healthy organizations call for overcoming "distrust" (Kwantes *et al.*, 2019), as it can lead to cynicism (Prajogo *et al.*, 2020). The research by Sen *et al.* (2021) emphasized reducing cynicism with the help of organizational support. Hence, to reduce the attrition and control the impact of cynicism in organizations (Randy Evans *et al.*, 2010), it becomes imperative to evoke citizenship behavior among employees and what can be a better way to enhance OCB than focusing on the key priority area of employees, "their career." Arguably, CMP adopted by the organizations have a significant association with organizational commitment, and the commitment is espoused to be the antecedent of organizational citizenship behavior (OCB). Studies reveal that citizenship behavior eventually leads to a reduction in turnover intention.

The focus on career management can enhance OCB and bring about effective organizational changes (Al-Shawabkeh, 2017). Despite research on career management, citizenship behavior and employee orientation separately, the importance of career management of the employee with supposedly "new career orientation" in enhancing the OCB has received scant attention. There is a need for unearthing the mechanism through which CMP impacts the OCB among "protean cohorts" across the genders. Since the present era is witnessing a "new career" orientation, it becomes imperative to investigate the interlinkages among CO, CMP and OCB to produce the desired outcomes.

Thus, keeping this background into focus, this study aims to address two aspects: First, the mediational role of CMP in the relationship between CO and OCB is discussed. It explains how the employees with new CO can trigger the role of CMP, and consequently, CMP can help them enact OCB. Secondly, how the gender of employees moderates the mediated impact of CMP in the relationship between CO and OCB is studied. It explains how the mediated relationship varies across genders. The study's novelty lies in exploring the moderated mediation analysis that has not been studied so far. The study can be instrumental in conducting future research in understanding the career management aspects of employees for the organization. The research has both theoretical and practical implications. Therefore, it can be a valuable contribution to the most recent literature available in career management and OCB. Also, the study results would have lasting and pragmatic implications for the organizations in lower attrition, enhanced performance and increased efficiency.

This research is divided into six parts: Section 2 discusses the literature. Section 3 is about research methodology. Section 4 discusses the analysis of the data, and Section 5 discusses the findings of the study. Section 6 discusses the theoretical, practical and policy implications of the study.

2. Literature review

2.1 Career orientation in the era of "new periodization"

To cope with the changing nature of career and behavior needs a mutual understanding between the employers and employees (Weng *et al.*, 2010; Weng and McElroy, 2012). The changes in technology, the revolution in communication and the advancement of technology have brought about changes in job design (Burke and Ng, 2006).

Consequently, employees try to empower their career escalation (Direnzo and Greenhaus, 2011) by setting the path by themselves (Weng and McElroy, 2012; Weng *et al.*, 2010) rather than relying on their employing organizations and getting benefits.

There are two essential characteristics of protean careerists, namely, being value-driven and being self-directed. Being self-directed relates to the intensity with which an individual takes strategic control of their career (Mirvis and Hall, 1994; Briscoe and Hall, 2006). At the same time, value-driven relates to the individual's consciousness toward their priorities

(Hall and Mirvis, 1996). A protean careerist is always the one who is open to exploring new possibilities and seeing his career as a series of the learning cycle, and (Hall, 2004) showing a calling attitude at the organization (Hall and Mirvis, 1995; Harrington and Hall, 2007; Hall and Chandler, 2005). While the protean career concept has gained due recognition (Hermann *et al.*, 2015; Lee *et al.*, 2014), empirical investigations continue to delve deeper into its psychological implications.

To cope with the present-day uncertainties of employment, career is increasingly transitioning from traditional ones to the protean (Arthur and Rousseau, 1996; Hall, 2002; Sullivan and Arthur, 2006). As proposed by the researchers, some of the attributes are positive interpersonal attitude (Seibert and Kraimer, 2001), self-directedness (Briscoe and Hall, 2006; Hall and Chandler, 2005), internal motivations (Gagne *et al.*, 1997); proactiveness (Briscoe *et al.*, 2006; Eby *et al.*, 2003), positive self-image (Judge and Bono, 2001), and adaptability and resilience (Lounsbury *et al.*, 2003), openness to career opportunities (Chiaburu *et al.*, 2006).

2.2 Organizational citizenship behavior (OCB)

OCB refers to the "behavior that is not formally rewarded but can be functional to the organization's operations (Smith *et al.*, 1983). Employees view OCB as a means of essentially compensating the organization in place of the actual fulfillment of the terms stated in the psychological (implicit) contract between employees and employers (Organ, 1988). The literature suggests that OCB can positively affect the success of the organizations through improvements in performance stability and the organizations' ability to environmental adaptations (Podsakoff *et al.*, 2000). Organ (1988) identified five dimensions of OCB. The dimensions are altruism, sportsmanship, conscientiousness, civic virtue and courtesy. Podsakoff *et al.* (2000) and Salehzadeh *et al.* (2015) further supported what Organ had stated. Ever since then, several dimensions are added to OCB. Sportsmanship, civic virtue, and conscientiousness among the employees can easily be differentiated. In contrast, one cannot distinguish between altruism and courtesy that easily.

There are two dimensions in which OCB was primarily organized, namely, altruism and compliance (Bateman and Organ, 1983). Later courtesy, conscientiousness, sportsmanship, altruism and civic virtues were added. The concept of OCB was proposed in yesteryears and studied in the US (Organ, 1988; Podsakoff *et al.*, 1990; Smith *et al.*, 1983). Studies have found several dimensions of OCB (Van Dyne *et al.*, 1994; Williams and Anderson, 1991).

2.3 Social exchange theory and citizenship behavior

The social exchange theory suggests that an individual's behavior is the result of an exchange process that occurs in organizations. This theory postulates that employees typify and humanize the organization by cultivating an exchange relationship that varies in attitudinal and behavioral reactions (Eisenberg *et al.*, 1997). Research by Saks (2006) highlighted the association of perceived organizational support with increased engagement at work, an antecedent to satisfaction, turnover intent and citizenship behavior.

2.4 Career management practices (CMPs)

Organizational career management (OCM) has been defined as a set of activities that companies sustain their employees' career development (Baruch and Peiperl, 2000). It helps them obtain promotions and pay raises and assist their transition into leadership positions (Vinkenburg and Weber, 2012) and support their employees' career (Kong *et al.*, 2012; Budhwar and Baruch, 2003). OCM is also known as an organization's support for career development or organizational sponsorship. It refers to the set of activities and the support by

the organization to take care of their employees' success in their respective careers (Ng *et al.*, 2016; Orpen, 1994). To enable OCM, several CMPs are followed in organizations. They are bundled into clusters (De Vos and Soens, 2008; De Vos and Cambre, 2017), which relate to various characteristics of the organization such as age, culture and size (Baruch and Peiperl, 2000). There are CMP defined by a group of authors (Eby *et al.*, 2003). The CMPs studied in the present study are development, potential assessment and feedback.

2.5 The rationale for the study

Evoking the OCB among employees can contribute tremendously to organizations. It has been found that OCB contributes toward the better performance of employees (Podsakoff *et al.*, 2000; Yen *et al.*, 2008), organizational effectiveness (Podsakoff and Mac Kenzie, 1997) and efficiency (Bateman and Organ, 1983). Employees view OCB as a means of essentially compensating the organization in place of the actual fulfillment of terms stated in the psychological (implicit) contract between employees and employers (Organ, 1988). When the employees start feeling that the organization meets the exchange obligations to attain career growth, they try to reciprocate through their acts and improve their performance (Organ and Konvsky, 1989). The emergent literature shows how Western countries observe a radical change in meanings attached to work and how the economic fluctuations have led to the change in the employer-employee psychological contract, which has consequently led to lost organizational commitment (Cappelli, 1999). Organizations are largely facing issues concerning citizenship behavior among protean cohorts. This calls for a need to unearth the mechanism through which OCB among the cohorts is regained. Thus, the research was conceptualized.

2.6 Gaps in the literature

To increase the commitment among employees, several CMPs are followed in the organizations. They are bundled (De Vos and Soens, 2008; De Vos and Cambre, 2017) into clusters related to various organizational characteristics such as age, culture and size (Baruch and Peiperl, 2000). Various research papers highlight the impact of career management on organizational commitment, which antecedes OCB. However, there is a dearth of research, particularly analyzing the mediational role of CMPs in the relationship between CO and OCB among the cohorts based on their gender. This gap in the emergent literature necessitates this study. This paper advances the extant literature on the said relationship. The study can be highly beneficial to provide impetus to the organizations with the employees having protean career orientation, focusing on CMPs to evoke OCB among employees, and increasing their productivity, efficiency, effectiveness and performance at work.

2.7 Theoretical foundations and hypotheses development

2.7.1 The link between new career orientation and career management practices. Today's employees derive satisfaction and benefit from the goals set for themselves (Weng *et al.*, 2010). Career goal influences the level of psychological well-being of individuals (Rahim and Siti-Rohaida, 2015). Kaminsky and Behrend (2015) examined the extent to which calling affected the career attitude of employees. Understanding the "calling" of employees has called for organizations to rethink the CMPs for their employees. A study by Holtschlag *et al.* (2020) demonstrated the indirect effect of CO to bring down the turnover intent of employees in the presence of appropriate CMP. The study, among others, sets the tone for positing that employees with the new CO would have a significant association with CMPs. Therefore, the following hypothesis was postulated:

H1. Protean career has a significant impact on career management practices.

2.7.2 Linking career management practices and citizenship behavior. A study by Lee *et al.* (2014) came up with a model by integrating the social cognitive model with well-being. Organizations are shifting their focus towards career management of employees, intending to enhance the satisfaction among employees. The role of career management has been conclusively proved to generate positive outcomes (Setyaningrum *et al.*, 2021). It is evident from the studies that individuals who have a more robust vision in terms of attainment during their career are likely to be more successful in their career (De Vos and Soens, 2008). A study by Chepkenoi (2021) further highlighted the role of a career in enhancing OCB. Therefore, if this consideration is held by the organization and good CMPs are employed for the employees, they are likely to go beyond the call of duty and exhibit organizational citizenship behavior. This gives way to the following hypothesis:

H2. Career management practices significantly impact organizational citizenship behavior.

2.7.3 Linking new career and citizenship behavior. A protean career attitude is often characterized by self-directed growth and a value-driven approach. Self-directed growth refers to "the degree to which an individual takes control of his/her career strategically" (Briscoe and Hall, 2006; Mirvis and Hall, 1994). A value-driven career approach refers to "the high consciousness of individual towards his or her priorities and is used as the standard in making and assessing decisions" (Hall and Mirvis, 1996). Grant and Ashford (2008) espouse that employees with such traits are very well aware of their achievements and are concerned with organizational preferences and priorities. Sullivan and Arthur (2006) suggested that employees who pursue this attitude are committed to their goals and values. Several studies espouse that the models of protean career indicate the declining commitment of employees toward the organization (Zaleska and de Menezes, 2007).

In contrast, Bolino *et al.* (2016) suggested the enactment of OCB being driven by "impression management." Since organizational commitment is the crucial determinant of loyalty, employee performance and turnover (Mowday *et al.*, 1982; Mathieu and Zajac, 1990; Shore and Wayne, 1993), the study would help us understand the behavioral consequences of career attitude. Hence the following hypotheses are proposed:

H3. Career orientation significantly impacts organizational citizenship behavior.

A change in the mindset of employees in the form of "new career" orientation calls for the organizations to wake up and rise for career management of the employees. It is evident from the studies highlighting the positive association between the perceived organizational support in terms of employee development and OCB (Jehanzeb, 2020). Substantial evidence suggests a link between the career practices followed in the organization and essential work outcomes (Schnake *et al.*, 2007; Chepkemoi, 2021). In the era characterized by new career orientation, wherein employees have a high urge for personal career management, employees are likely to enact OCB (to reciprocate the behavior) due to CMPs adopted by the organization. Therefore, it becomes imperative to study how CMPs can mediate the relationship between CO and OCB. Thus, the following hypothesis is postulated:

H4. CMP significantly mediates the relationship between career orientation and citizenship behavior.

The gender differences in the organization arise due to the difference in responsibilities, degree of experience and stress-bearing ability that the genders encounter differently (Petrides and Furnham, 2006). Both the genders experience different roles both in their jobs and in life (Colley and Comber, 2003). The varied nature of employment and occupations makes individuals look for their choices in terms of their careers (Yildis *et al.*, 2015). The priorities are subjected to one's gender. Thus, to gauge the difference in the CMP-mediated

relationship of CO and OCB that may arise because of gender, the following hypothesis was postulated.

H5. Gender significantly moderates the CMP-mediated relationship between career orientation and citizenship behavior.

3. Research methodology

3.1 Participants and procedure

The study comprised the respondents belonging to middle-level management positions across diverse industries in Delhi and NCR. An online questionnaire was created for collecting the information from the employees. The demographic data were taken from the respondents in the first part of the questionnaire. The second part comprises the questions about career orientation, organizational citizenship behavior, and CMPs as the primary constructs. The sub-constructs of OCB were altruism, sportsmanship, conscientiousness, courtesy, civic virtues, and the sub-constructs of CO were self-directed growth and value-driven career.

We used a five-point Likert scale with a range of 1–5 to measure the responses. On the scale, one meant strongly disagree, and five meant strongly agree. The order of the questions in the OCB inventory was also changed, keeping in mind the essence of avoiding sequential response bias (Esfandiar *et al.*, 2019).

The pilot test was conducted before distributing and circulating the questionnaire on a full scale. We collected fifty responses during the pilot testing phase. The Cronbach's alpha score came out to be more than 0.75 for all the questionnaire items, and thus, it concluded that the scale was internally consistent and could be used for conducting the study (Hair *et al.*, 2010). After the pilot study, the questionnaires were circulated online among the respondents for complete research.

3.2 Construct measurement

The constructs used in the study have been explained in Section 3.2.1 to 3.2.3.

3.2.1 Organizational citizenship behavior. The OCB framework, as proposed by Organ (1988), has been consistently treated over a good number of studies (Le Pine *et al.*, 2002). For capturing OCB, a 24 item inventory by Podsakoff *et al.* (2000) was used in the study since it is easy to understand and has been widely used in researches (particularly in the context of Asian countries). The sub-constructs of OCB were altruism, sportsmanship, conscientiousness, courtesy and civic virtues. A sample item is, "I believe in giving an honest day's work for an honest day's pay."

3.2.2 Career management practices. For the CMPs, three dimensions of career management strategies were considered, namely development of employees, feedback, and potential assessment (Eby *et al.*, 2003). A sample item is, "I feel that the organization provides me with ample opportunity to develop my career."

3.2.3 Career orientation of employees. For capturing the responses of the sample population concerning career orientation, the popular scale by Briscoe and Hall (2006) was used. A sample item is, "When development opportunities have not been offered by my company, I've sought them out on my own."

3.3 Sampling and data collection

The framework for conducting the sample study was non-probabilistic, as the apt sampling frame is generally not available for such kind of study. A convenience sampling method was used. This act is entirely in line with the studies conducted in the recent past. A self-

administered questionnaire was used to collect the data for the analysis. The duration of these data collection was fixed as four months. Of 300 questionnaires that were circulated, 39 were returned incomplete, and 47 questionnaires were not filled. Hence, a total of 214 useable responses were considered for the study. The overall response rate was 71.33, which is on the higher side. Malhotra and Grover (1998) emphasized that the response rate of 20% was enough for the positive administration of the survey.

3.4 Data analysis

We used the structural equation modeling (SEM) approach to study the hypothesized relationship between the various constructs since SEM is considered an appropriate method for multivariate analysis. AMOS 25 was used for checking the associations and impacts. First, the model's fitness was assessed using measurement indices, followed by a reliability and validity check of the constructs. After it, the mediation analysis was checked using AMOS for both the genders.

4. Results and findings

This section discusses the demographic profile and descriptive statistics (mean, standard deviation, standard error and internal consistencies of the constructs); measurement model for the study; convergent and discriminant validity; the structural model used for the study with the results of Hypothesis 1 through Hypothesis 3; the mediational influence of CMP in the CO-OCB relationship and explanations of moderated mediation analysis with the results of Hypothesis 4 and 5.

4.1 Demographic profile and descriptive statistics of the respondents

The study respondents were employed across diverse industries in Delhi, NCR of India. 55% were males, and 45% were females. 38% of employees were between 20–30 years of age; 36% were between 30–40 years, and 26% were between 40 and 50 years old. 16% of employees possessed less than a year's experience; 30% of them had 1–2 years of experience; 33% had 2–5 years of experience, and 21% had more than five years' experience.

There are three constructs used in the study, namely COCO, career management strategies (CMP) and OCB. Table 1 presents the internal consistencies, mean, standard error and standard deviation of each of the constructs used in the study. Moreover, the correlation among the constructs is also given in the table. The correlation between the scales used in the study shows the direction of association among them, with all the relationships significant at $p < 0.05$. Tabachnick *et al.* (2007) suggest that the correlation coefficient is less than 0.7, representing no chances of multicollinearity in regression. In our study, however, the inter-construct correlation is less than 0.5; hence, the measures were considered appropriate for further analysis.

Table 1. Descriptive statistics, correlations, mean and standard deviation

Factors used in the study	Internal consistencies	Mean	Standard error	Standard deviation	1	2	3
1. CO	0.702	47.47	0.39	6.96			
2. CMP	0.771	11.08	0.15	2.70	*0.48		
3. OCB	0.840	91.95	0.57	10.04	*0.37	*0.45	

Note(s): $n = 214$; *$p < 0.01$; career orientation (CO); career management practices (CMP) and organizational citizenship behavior (OCB)

4.2 Measurement model for the study

A two-fold approach was used for testing the hypothesis. First, the measurement model was tested using the indices (Anderson and Gerbing, 1988). Second, we checked the reliability and validity of the model. We followed Podsakoff *et al.* (2003)'s recommendations to ensure that it controls the influence of common method bias in the study. It suggested using established scales only. The explanation of procedures to our participants was also very well explained at the outset, and they were guaranteed anonymity. Moreover, we did the psychological separations between the variables by adding some filler items to the scale. The confirmatory factor analysis (CFA) was carried out on the data set to check if our model fits the data. For testing the model fit, five indicators are used, namely cmin/df; GFI; RMSEA; CFI and AGFI. Normed chi-square (cmin/df) measures chi-square statistics and is obtained by dividing the chi-square by degree of freedom. Hair *et al.* (2010) assert that this measure overcomes some of the limitations of chi-square and thus handles the problem of model complexity. It should range between one and three. The goodness of fit (GFI) is a measure to test the fitness of the model. Tanaka and Huba (1989) espouse that GFI is analogous to r square in regression analysis. The value of GFI should be between zero and one, with a higher value demonstrating a good fit. The root-mean-square of approximation (RMSEA) is a population-based index and is less sensitive to the sample size. This measure is used to correct the limitations of chi-square GFI and is one of the most widely used measures. The lesser the value of RMSEA, the better is the model fit. Hair *et al.* (2010) define the cut-off value as 0.05. The comparative fit index (CFI) is an improved version of the normed fit index, and its value should range between zero and one, with higher values demonstrating a better fit. The threshold limits of indices used for checking the fitness of the model are chi-square divided by degree of freedom (cmin/df $\leq$ 3); goodness of fit index (GFI $\geq$ 0.90); adjusted goodness of fit (AGFI > 0.85); RMSEA (<0.05); comparative fit index (CFI $\geq$ 0.90) as suggested by Hair *et al.* (2010). The CFA results showed a good model fit across both genders. (Males: cmin/df = 1.18; GFI = 0.910; AGFI = 0.890; RMSEA = 0.042; CFI = 0.923 and for female group: cmin/df = 1.46; GFI = 0.920; AGFI = 0.890; RMSEA = 0.043; CFI = 0.963). Since the values of test indices for CFA were well within the accepted range suggested by Hair *et al.* (2010), the model demonstrated a clear distinctiveness of the measures and was considered fit.

4.3 Convergent validity and discriminant validity

The model's validity was assessed using convergent validity assessments and discriminant validity checks (Garver and Mentzer, 1999). Table 2 shows the reliability and validity for the male group, and Table 3 shows the reliability and validity analysis for the female groups. It explains CR for all the constructs used in the study. Convergent validity for all the variables is higher than 0.7, and factor loading is more than 0.5, which is the cut-off range. It shows that the variables are unidimensional and are reliable for the study (Hair *et al.*, 2010).

4.4 Structural model

We used regression analysis to test hypothesis 1 to hypothesis 3. The results of regression analysis have been presented in Table 4. It shows the beta value, standard error, critical ratio and p-value.

Hypothesis 1 discusses the direct effect of CMP on OCB. Since the beta value (0.323) for CO impacting CMP is significant, hence Hypothesis 1 stating that CO significantly impacts CMP is supported. Hence, our argument that employees' new CO is likely to perceive the CMP well is supported. Hypothesis 2 argues that CMPs are reasonable to make employees enact OCB. Since the beta value (0.398) of CMP impacting OCB is significant, hypothesis 2 is supported. Hypothesis 3 posits that employees with new CO are likely to engage in OCB. The findings in Table 4 confirm the hypothesis since the beta value (0.281) of CO impacting OCB is statistically significant; hence, we can conclude that hypothesis 3 is supported.

	CR	AVE	MSV	sd	vd	Alt	spo	civic	court	con
Ps	0.851	0.557	0.667	*0.676*						
pv	0.766	0.594	0.667	0.816***	*0.542*					
oa	0.840	0.635	0.226	0.451***	0.348***	*0.660*				
os	0.702	0.626	0.084	−0.152†	−0.282**	0.105	*0.571*			
ov	0.799	0.586	0.567	0.176†	0.220*	0.306**	−0.29**	*0.620*		
ocu	0.701	0.622	0.897	0.468***	0.439***	0.476***	−0.179*	0.692***	*0.568*	
oc	0.706	0.524	0.897	0.378***	0.388***	0.465***	−0.28**	0.753***	0.947***	*0.569*

Note(s): Composite reliability (CR); average variance extracted (AVE); maximum shared variance (MSV); self-directed growth (ps); value driven (pv); altruism (alt); sportsmanship (spo); civic virtue (civic) and conscientiousness (cons). †$p < 0.100$, *$p < 0.050$, **$p < 0.010$, ***$p < 0.001$

Table 2.
Reliability and validity analysis of the male group

Table 3.
Reliability and validity analysis of the female group

	CR	AVE	MSV	sd	vd	alt	spo	civic	court	con
ps	0.800	0.588	0.324	0.623						
pv	0.780	0.518	0.260	0.476**	0.564					
oa	0.837	0.731	0.324	0.569***	0.510**	0.657				
os	0.738	0.678	0.081	−0.024	−0.007	0.285*	0.615			
ov	0.718	0.734	0.556	0.265*	0.198	0.435**	−0.012	0.628		
ocu	0.785	0.615	0.306	0.367**	0.286†	0.512**	0.211	0.746***	0.561	
oc	0.772	0.706	0.516	0.165	0.223	0.542***	0.250†	0.689***	0.952***	0.637

Note(s): Composite reliability (CR); average variance extracted (AVE); maximum shared variance (MSV); self-directed growth (ps); value-driven (pv); altruism (alt); sportsmanship (spo); civic virtue (civic) and conscientiousness (cons). †$p < 0.100$, *$p < 0.050$, **$p < 0.010$, ***$p < 0.001$

4.5 Mediational influence of CMP in the CO-OCB relationship: explanation of moderated mediation

To test hypotheses 4 and 5, we used mediation analysis using AMOS software. The moderated mediation was tested using the steps laid by Preacher *et al.* (2007). The independent variables were first standardized so that tests of interaction could be carried out, and the dependent variables were unstandardized (Aiken and West, 1991)

For testing the mediation role of CMP in the CO-OCB relationship, we followed the steps outlined by Baron and Kenny (1986). The conditions for mediation analysis comprise the correlation of the independent variable with the outcome variable. The independent variable must be significantly associated with the mediator variable, and the mediator variable must be significantly correlated with the outcome variable. Moreover, when the independent and mediator variables are entered together, the independent variable should lead to a drop in strength and/or significance (Alfez *et al.*, 2013). The conditions for mediation analysis were met, and after that, we performed the mediation analysis.

We checked the mediation influence of CMPs in the relationship between CO and OCB among the groups based on gender. The mediation analysis results have been presented in Table 5. Table 5 shows the direct effect, indirect effect, total effect, standard error, ULCI and LLCI of the mediation analysis used in the study. It is evident from Table 3 that CO has a significant impact on OCB, and CMP has a significant impact on OCB. The two subgroups were males and females. For the study, the confidence interval was bootstrapped 5,000 times.

Hypothesis	Relationship	beta value	SE	CR	*p*-value	Supported/Not supported
1	CO → CMP	0.323	0.021	5.94	***	Supported
2	CMP → OCB	0.398	0.195	7.57	***	Supported
3	CO → OCB	0.281	0.079	5.11	***	Supported

Note(s): Critical ratio (CR); standard error (S.E); career orientation (CO); career management practices (CMP) and organizational citizenship nehavior (OCB)

Table 4. Result of structural analysis

CO→CMP→OCB	Beta	S.E	*p* value	LLCI	ULCI
Study sample A (males)					
CO → OCB	0.196	0.075	0.003	0.085	0.334
CO → CMP	0.306	0.079	0.001	0.171	0.436
CMP → OCB	0.257	0.066	0.002	0.137	0.354
Total effect of CO → OCB/CMP	0.275	0.071	0.001	0.162	0.399
Study sample B (females)					
CO → OCB	0.124	0.164	0.415	−0.127	0.420
CO → CMP	0.376	0.118	0.004	0.169	0.561
CMP → OCB	0.535	0.133	0.002	0.291	0.737
Total effect of CO→OCB/CMP	0.325	0.14	0.016	0.098	0.568
Indirect effect					
CO → CMP → OCB (males)	0.079	0.028	0.001	0.041	0.134
CO → CMP → OCB (females)	0.201	0.085	0.002	0.086	0.370

Note(s): Critical ratio (CR); standard error (S.E); career orientation (CO); career management practices (CMP); organizational citizenship behavior (OCB); lower limit confidence Interval (LLCI) and upper limit confidence limit (ULCI)

Table 5. Mediation analysis results for the study (direct, indirect and total effects)

Using the bootstrapping technique, we found the partial mediation role of CMP in the CO-OCB relationship among male employees. According to bootstrap results, upon the addition of CMP as a mediator, the main effect of CO on OCB (Sample1 (Males): Beta = 0.196; LLCI = 0.085; ULCI = 0.334; $p \leq 0.05$) remained significant for the males' subgroup (as presented in Table 5). But there was a drop in the beta value from 0.281 (as shown in Table 4) to .196, with the p-value remaining significant. This makes it evident that the CMP partially mediated the relationship of CO and OCB in male employees. The effect is significant because there is no zero between ULCI and LLCI, providing additional support for partial mediation.

However, in the case of the female sub-group, the addition of CMP as a mediator caused the relationship to become insignificant (Beta = 0.124; LLCI = 0.415; ULCI = −0.127; $p = 0.420$). Therefore, we can infer that for the female subgroup, CMP fully mediated the relationship between CO and OCB. The effect is significant because there is no zero between ULCI and LLCI, providing additional support for full mediation.

Our results show that CMPs are essentially imperative in the sustenance of OCB among the employees with new career orientation, who would otherwise disengage at work. In other words, the new CO employees' OCB is shaped by perceived CMPs in both genders.

5. Discussion

The central conjuncture of the study revolves around understanding the mediating and moderating processes through which the CO of employees is linked with citizenship behavior. Drawing on the social exchange theory, the model posited the mediation role of CMP in the CO-OCB relationship and further investigated if the gender of employees moderates the association between CO-OCB. Data from employees working at middle-level management positions in Delhi and NCR of India support our model. The study has theoretical, practical and policy implications.

This study focuses light on the nature of the relationship between CMPs. It is seen that where the employees had a positive perception of career management initiatives, the OCB was high. This study adds to the growing and extant literature examining the effects of HRM practices on citizenship behavior (Kuvaas, 2008; Nishii *et al.*, 2008).

Hypothesis 1 discusses the direct effect of CMP on OCB. Since the beta value (0.323) for CO impacting CMP is significant, hence, Hypothesis 1 stating that CO significantly impacts CMP is supported. Hence, our argument that employees' new CO is likely to perceive the CMP well is supported. Hypothesis 2 argues that CMPs are likely to make employees enact OCB. Since the beta value (0.398) of CMP impacting OCB is significant, hypothesis 2 is supported. Hypothesis 3 posits that employees with new CO are likely to engage in OCB. The findings in Table 4 confirm the hypothesis since the beta value (0.281) of CO impacting OCB is statistically significant; hence, we can conclude that hypothesis 3 is supported.

Hypothesis 4 posited that CMP significantly mediates the relationship between CO and citizenship behavior. To test this hypothesis, mediation analysis was used. The results of the mediation analysis support the hypothesis. Our results show that CMPs are essentially imperative in the sustenance of OCB among the employees with new career orientation. In other words, the new CO employees' OCB is shaped by perceived CMPs.

Further, Hypothesis 5 posited that gender significantly moderates the CMP-mediated relationship between CO and citizenship behavior. To check this hypothesis, the two subgroups formed were males and females, and the confidence interval was bootstrapped 5,000 times. The results show that for the female sub-group, CMP fully mediated the relationship between CO and OCB. In contrast, for male subgroups, CMP partially mediated the relationship between CO and OCB.

Our study illuminates CMPs as a catalyst for the association between CO and organizational citizenship behavior. The employees who perceive CMPs adopted by the organization positively are more likely to exhibit and enact citizenship behavior. They are

likely to go beyond the call of duty. OCB has far-reaching implications. Introducing a new mediator in the relationship between CO and OCB added new perspectives to the existing literature, which primarily focused on commitment, organizational support and employee satisfaction (Conway and Monks, 2009; Snape and Redman, 2010). The changing CO of employees places high value on self-driven growth and value-driven career. We argue that these shifting orientations eventually compel the organizations to rethink the CMPs they have been following. The perception of employees concerning CMPs drives their citizenship behavior. This study considers the perception of employees concerning feedback, potential assessment and development as CMPs. There is immense scope for further research in the area.

Finally, our research highlights that gender is a moderating variable in the CMP-mediated relationship between CO and OCB. The results suggest that CMP fully mediated the relationship between CO and OCB in the case of female employees. Among male employees, CMPs can partly enact the link between CO and OCB. Therefore, we can say that the extent to which CMP mediates the relationship between CO-OCB is contingent upon the gender of employees.

6. Theoretical, practical and policy implications

This study adds to the growing and extant literature by examining the effects of HRM practices on citizenship behavior (Kuvaas, 2008; Nishii *et al.*, 2008). First, the contribution is in terms of inclusion of affective state in gaining an insight into the impact of CMP as a mediator in the relationship between predictor and outcome variable. Our research bridges the two consistently independent researches into a coherent model. As posited, we found a significant mediation effect of CMP, which clearly states that introducing CMPs can have multi-fold benefits among protean careerists. Also, CMPs can be instrumental in the enactment of citizenship behavior among employees. Thus, to strengthen the current understanding of career and its impact on OCB, the study incorporated the concept from social exchange theory and can be used for future studies in career management. It sets empirical investigation of several strands of career, cynicism and work behavior.

Second, the investigation of gender as a moderator in the CMP-mediated relationship is a significant contribution. No such study wherein moderated mediation in such a relationship is established has been found. Previous studies highlighted the association of CMP with CO (Volmer and Spurk, 2011), CO with OCB (Presti *et al.*, 2019) and CMP with OCB (Al-Shawabkeh, 2017). However, our study is the first to conceptualize moderation in the mediation analysis *per se*. We found that CMP reduced the effect of CO on OCB from a beta value of 0.28 to 0.196 (for male employees) but remained significant, making it a case of partial mediation (Baron and Kenny, 1986). However, in the case of women, the effect lowered down and became insignificant, stating that it is the case of full mediation. The research findings correspond to the social exchange theory (Cook *et al.*, 2013), which postulates that employees typify and humanize the organization by cultivating an exchange relationship that varies in attitudinal and behavioral reactions (Eisenberg *et al.*, 1997).

The study also has some pragmatic implications. In times of changes and economic downturns, employees are essential for a competitive edge (Costa and Neves, 2017). With the employees possessing the new career orientation, OCB enactment becomes a must. In the previous studies, OCB is well studied as an antecedent to reduced absenteeism (Khalid and Ali, 2005) (Chughtai and Zafar, 2006), reduced turnover and employee satisfaction. Our study can be instrumental in the enactment of OCB through CMP among the protean cohort. Organizations can develop effective career management strategies to improve work behavior and productivity by evoking OCB among employees. Career strategists can use it to understand the role of gender in administering the outcomes in the form of the correct CMPs for men and women employees.

Taking into consideration the gender aspects, organizations can devise the customized strategies for men and women. This research further develops the base for understanding the specific techniques that work for groups based on gender and opens the doors for future research for tracing the practices apart from CMP to reinforce the enactment of OCB among employees further.

The study can be instrumental for the policymakers in the organizations. They can take cues from empirical evidence and understand the central proposition of the social exchange theory. The theory suggests that organizations that can cultivate the practice of reciprocity can elicit positive responses in terms of behavior and attitude among employees. Our findings are also consistent with the theory. Thus, this study can be used to effectively enhance the psychological contracts between the employer and employees by developing the appropriate policies across the genders. It is very clear from the study that career management is instrumental for enacting OCB among cohorts, and they do consider CMP within the context of their environment. Therefore, career management contribution helps create a holistic and employee-centric approach to help employees enact OCB.

References

Aiken, L.S., West, S.G. and Reno, R.R. (1991), *Multiple Regression: Testing and Interpreting Interactions*, Sage, Newbury Park.

Alfez, K., Shantz, A.D., Truss, C. and Soane, E.C. (2013), "The link between perceived human resource management practices, engagement and employee behavior: a moderated mediation model", *The International Journal of Human Resource Management*, Vol. 24 No. 2, pp. 330-351, doi: 10.1080/09585192.2012.679950.

Al-Shawabkeh, K.M. (2017), "Career path development and its impact on organizational citizenship behavior in greater Amman municipality", *International Journal of Business and Management*, Vol. 12 No. 3, p. 79.

Anderson, J.C. and Gerbing, D.W. (1988), "Structural equation modeling in practice: a review and recommended two-step approach: psychological bulletin", Vol. 103 No. 3, pp. 411-423.

Arthur, M.B. and Rousseau, D.M. (1996), "A career lexicon for the 21st century", *Academy of Management Perspectives*, Vol. 10 No. 4, pp. 28-39.

Baron, R.M. and Kenny, D.A. (1986), "The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations", *Journal of Personality and Social Psychology*, Vol. 52 No. 6, p. 1173.

Baruch, Y. and Piperl, M. (2000), "Career management practices: an empirical survey and implications", *Human Resource Management*, Vol. 39 No. 4, pp. 347-366.

Bateman, T.S. and Organ, D.W. (1983), "Job satisfaction and the good soldier: the relationship between affect and employee citizenship", *Academy of Management Journal*, Vol. 26 No. 4, pp. 587-595.

Bolino, M., Long, D. and Turnley, W. (2016), "Impression management in organizations: critical questions, answers, and areas for future research", *Annual Review of Organizational Psychology and Organizational Behavior*, Vol. 3, pp. 377-406.

Briscoe, J.P. and Hall, D.T. (2006), "The interplay of boundaryless and protean careers: combinations and implications", *Journal of Vocational Behaviour*, Vol. 69 No. 1, pp. 4-18.

Briscoe, J.P., Hall, D.T. and DeMuth, R.L.F. (2006), "Protean and boundaryless careers: an empirical exploration", *Journal of Vocational Behaviour*, Vol. 69 No. 1, pp. 30-47.

Budhwar, P.S. and Baruch, Y. (2003), "Career management practices in India: an empirical study", *International Journal of Manpower*, Vol. 24 No. 6, pp. 699-719.

Burke, R.J. and Ng, E. (2006), "The changing nature of work and organizations: Implications for human resource management", *Human Resource Management Review*, Vol. 16 No. 2, pp. 86-94.

Cappelli, P. (1999), "Career jobs are dead", *California Management Review*, Vol. 42 No. 1, pp. 146-167.

Chepkemoi, J. (2021), *Employee Retention Strategies, Career Development Practices, and Organizational Citizenship Behavior in Environmental Agencies within North Rift Region*, Doctoral dissertation, Moi University, Kenya.

Chiaburu, D.S., Baker, V.L. and Pitariu, A.H. (2006), "Beyond being proactive: what (else) matters for career self-management behaviours?", *Career Development International*, Vol. 11 No. 7, pp. 619-632.

Chughtai, A.A. and Zafar, S. (2006), "Antecedents and consequences of organizational commitment among Pakistani university teachers", *Applied HRM Research*, Vol. 11 No. 1, p. 39.

Colley, A. and Comber, C. (2003), "Age and gender differences in computer use and attitudes among secondary school students: what has changed?", *Educational Research*, Vol. 45 No. 2, pp. 155-165.

Conway, E. and Monks, K. (2009), "Unravelling the complexities of high commitment: an employee-level analysis", *Human Resource Management Journal*, Vol. 19 No. 2, pp. 140-158.

Cook, K.S., Cheshire, C., Rice, E.R. and Nakagawa, S. (2013), "Social exchange theory", *Handbook of Social Psychology*, Springer, pp. 61-88.

Costa, S. and Neves, P. (2017), "Job insecurity and work outcomes: the role of psychological contract breach and positive psychological capital", *Work and Stress*, Vol. 31 No. 4, pp. 375-394.

De Vos, A. and Cambré, B. (2017), "Career management in high-performing organizations: a set-theoretic approach", *Human Resource Management*, Vol. 56 No. 3, pp. 501-518.

De Vos, A. and Soens, N. (2008), "Protean attitude and career success: the mediating role of self-management", *Journal of Vocational Behavior*, Vol. 73 No. 3, pp. 449-456.

Direnzo, M.S. and Greenhaus, J.H. (2011), "Job search and voluntary turnover in a boundaryless world: a control theory perspective", *Academy of Management Review*, Vol. 36 No. 3, pp. 567-589.

Eby, L.T., Butts, M. and Lockwood, A. (2003), "Predictors of success in the era of the boundaryless career", *Journal of Organisational Behavior: The International Journal of Industrial, Occupational and Organisational Psychology and Behavior*, Vol. 24 No. 6, pp. 689-708.

Eisenberg, N., Fabes, R.A., Shepard, S.A., Murphy, B.C., Guthrie, I.K., Jones, S. and Maszk, P. (1997), "Contemporaneous and longitudinal prediction of children's social functioning from regulation and emotionality", *Child Development*, Vol. 68 No. 4, pp. 642-664.

Esfandiar, K., Sharifi-Tehrani, M., Pratt, S. and Altinay, L. (2019), "Understanding entrepreneurial intentions: a developed integrated structural model approach", *Journal of Business Research*, Vol. 94, pp. 172-182.

Gagné, M., Senecal, C.B. and Koestner, R. (1997), "Proximal job characteristics, feelings of empowerment, and intrinsic motivation: a multidimensional model 1", *Journal of Applied Social Psychology*, Vol. 27 No. 14, pp. 1222-1240.

Garver, M.S. and Mentzer, J.T. (1999), "Logistics research methods: employing structural equation modeling to test for construct validity", *Journal of Business Logistics*, Vol. 20 No. 1, p. 33.

Grant, A.M. and Ashford, S.J. (2008), "The dynamics of proactivity at work", *Research in Organisational Behaviour"*, Vol. 28, pp. 3-34.

Hair, J.F., Anderson, R.E., Babin, B.J. and Black, W.C. (2010), "Multivariate data analysis: a global perspective", Vol. 5 No. 6, pp. 629-686.

Hall, D.T. (2002), *Careers in and Out of Organisations*, Sage, Thousand Oaks.

Hall, D.T. (2004), "The protean career: a quarter-century journey", *Journal of Vocational Behavior*, Vol. 65 No. 1, pp. 1-13.

Hall, D.T. and Chandler, D.E. (2005), "Psychological success: when the career is a calling", *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, Vol. 26 No. 2, pp. 155-176.

Hall, D.T. and Mirvis, P.H. (1995), "The new career contract: developing the whole person at midlife and beyond", *Journal of Vocational Behavior*, Vol. 47 No. 3, pp. 269-289.

Hall, D.T. and Mirvis, P.H. (1996), "The new protean career: psychological success and the path with a heart", in Hall, D.T. and Associates (Eds), *The Career is Dead—Long Live the Career*, Jossey Bass, San Francisco, pp. 15-45.

Harrington, B. and Hall, D.T. (2007), *Career Management & Work-Life Integration: Using Self-Assessment to Navigate Contemporary Careers*, Sage.

Hermann, K.M., Kocisky, T., Grefenstette, E., Espeholt, L., Kay, W., Suleyman, M. and Blunsom, P. (2015), "Teaching machines to read and comprehend", in *Advances in Neural Information Processing Systems*, pp. 1693-1701.

Holtschlag, C., Masuda, A.D., Reiche, B.S. and Morales, C. (2020), "Why do millennials stay in their jobs? The roles of protean career orientation, goal progress and organizational career management", *Journal of Vocational Behavior*, Vol. 118, p. 103366.

Jehanzeb, K. (2020), "Does perceived organizational support and employee development influence organizational citizenship behavior?", *European Journal of Training and Development*, Vol. 44 Nos 6/7, pp. 637-657.

Judge, T.A. and Bono, J.E. (2001), "Relationship of core self-evaluations traits—self-esteem, generalized self-efficacy, locus of control, and emotional stability—with job satisfaction and job performance: a meta-analysis", *Journal of Applied Psychology*, Vol. 86 No. 1, pp. 80-92.

Kaminsky, S.E. and Behrend, T.S. (2015), "Career choice and calling: integrating calling and social cognitive career theory", *Journal of Career Assessment*, Vol. 23 No. 3, pp. 383-398.

Khalid, S.A. and Ali, H. (2005), "The effect of organizational citizenship behavior on withdrawal behavior: a Malaysian study", *International Journal of Management and Entrepreneurship*, Vol. 1 No. 1, pp. 30-40.

Kong, F., Zhao, J. and You, X. (2012), "Emotional intelligence and life satisfaction in Chinese university students: the mediating role of self-esteem and social support", *Personality and Individual Differences*, Vol. 53 No. 8, pp. 1039-1043.

Kuvaas, B. (2008), "An exploration of how the employee–organization relationship affects the linkage between perception of developmental human resource practices and employee outcomes", *Journal of Management Studies*, Vol. 45 No. 1, pp. 1-25.

Kwantes, C.T. and Bond, M.H. (2019), "Organizational justice and autonomy as moderators of the relationship between social and organizational cynicism", *Personality and Individual Differences*, Vol. 151, p. 109391.

Lee, C.I., Felps, W. and Baruch, Y. (2014), "Toward a taxonomy of career studies through bibliometric visualization", *Journal of Vocational Behavior*, Vol. 85 No. 3, pp. 339-351.

LePine, J.A., Erez, A. and Johnson, D.E. (2002), "The nature and dimensionality of organizational citizenship behavior: a critical review and meta-analysis", *Journal of Applied Psychology"*, Vol. 87 No. 1, p. 52.

Lounsbury, J.W., Loveland, J.M., Sundstrom, E.D., Gibson, L.W., Drost, A.W. and Hamrick, F.L. (2003), "An investigation of personality traits in relation to career satisfaction", *Journal of Career Assessment*, Vol. 11 No. 3, pp. 287-307.

Malhotra, M.K. and Grover, V. (1998), "An assessment of survey research in POM: from constructs to theory", *Journal of Operations Management"*, Vol. 16 No. 4, pp. 407-425.

Mathieu, J.E. and Zajac, D.M. (1990), "A review and meta-analysis of the antecedents correlate, and consequences of organisational commitment", *Psychological Bulletin*, Vol. 108 No. 2, p. 171.

Mirvis, P.H. and Hall, D.T. (1994), "Psychological success and the boundaryless career", *Journal of Organisational Behaviour*, Vol. 15 No. 4, pp. 365-380.

Mowday, R.T., Porter, L.W. and Steers, R. (1982), "Organisational linkages: the psychology of commitment, absenteeism, and turnover".

Ng, T.W., Lam, S.S. and Feldman, D.C. (2016), "Organizational citizenship behavior and counterproductive work behavior: do males and females differ?", *Journal of Vocational Behavior*, Vol. 93, pp. 11-32.

Nishii, L.H., Lepak, D.P. and Schneider, B. (2008), "Employee attributions of the 'why' of HR practices: their effects on employee attitudes and behaviors, and customer satisfaction", *Personnel Psychology*, Vol. 61 No. 3, pp. 503-545.

Organ, D.W. and Konvsky (1989), "Cognitive versus effective determinants of organization citizenship", *Journal of Applied Psychology*, Vol. 74, pp. 157-164.

Organ, D.W. (1988), *Organisational Citizenship Behaviour: the Good Soldier Syndrome*, Lexington Books/DC Heath and Com, Lexington.

Orpen, C. (1994), "The effects of organizational and individual career management on career success", *International Journal of Manpower*, Vol. 15 No. 1, pp. 27-37.

Petrides, K.V. and Furnham, A. (2006), "The role of trait emotional intelligence in a gender-specific model of organizational variables", *Journal of Applied Social Psychology*, Vol. 36 No. 2, pp. 552-569.

Podsakoff, P.M. and MacKenzie, S.B. (1997), "Impact of organisational citizenship behaviour on organisational performance: a review and suggestion for future research", *Human Performance*, Vol. 10 No. 2, pp. 133-151.

Podsakoff, P.M., MacKenzie, S.B., Moorman, R.H. and Fetter, R. (1990), "Transformational leader behaviours and their effects on followers' trust in leader, satisfaction, and organisational citizenship behaviours", *The Leadership Quarterly*, Vol. 1 No. 2, pp. 107-142.

Podsakoff, P.M., MacKenzie, S.B., Paine, J.B. and Bachrach, D.G. (2000), "Organisational citizenship behaviours: a critical review of the theoretical and empirical literature and suggestions for future research", *Journal of Management*, Vol. 26 No. 3, pp. 513-563.

Podsakoff, P.M., MacKenzie, S.B., Lee, J.Y. and Podsakoff, N.P. (2003), "Common method biases in behavioral research: a critical review of the literature and recommended remedies", *Journal of Applied Psychology*, Vol. 88 No. 5, p. 879.

Prajogo, W., Kusumawati, H. and Wijaya, N.H.S. (2020), "Organizational culture as determinant of employee-based antecedents to organizational cynicism", *Humanities and Social Sciences Reviews*, Vol. 8 No. 1, pp. 881-886.

Preacher, K.J., Rucker, D.D. and Hayes, A.F. (2007), "Addressing moderated mediation hypotheses: theory, methods, and prescriptions", *Multivariate Behavioral Research*, Vol. 42 No. 1, pp. 185-227.

Presti, A.L., Manuti, A. and Briscoe, J.P. (2019), "Organizational citizenship behaviors in the era of changing employment patterns: the complementary roles of psychological contracts and protean and boundaryless careers", *Career Development International*, Vol. 24 No. 2, pp. 127-145.

Rafiei, M. (2017), "The effect of emotional quotient on organizational citizenship behavior in some Iranian hospitals", *Kybernetes*, Vol. 46 No. 7, pp. 1189-1203, doi: 10.1108/K-01-2017-0001.

Rahim, N.B. and Siti-Rohaida, M.Z. (2015), "Protean COand career goal development: do they predict engineer's psychological well-being", *Procedia-Social and Behavioral Sciences*, Vol. 172, pp. 270-277.

Randy Evans, W., Goodman, J.M. and Davis, W.D. (2010), "The impact of perceived corporate citizenship on organizational cynicism, OCB, and employee deviance", *Human Performance*, Vol. 24 No. 1, pp. 79-97, doi: 10.1080/08959285.2010.530632.

Rodrigues, R., Guest, D. and Oliveira, T.C. (2021), "Are participants in the 'new career' good organizational citizens?", *The International Journal of Human Resource Management*, Vol. 1, p. 27.

Ronn, M. (2010), *The Relationship between Career Management and Organisational Commitment: the Moderating Effect of Openness to Experience*, Doctoral dissertation, University of Stellenbosch, Stellenbosch.

Saks, A.M. (2006), "Antecedents and consequences of employee engagement", *Journal of Managerial Psychology*, Vol. 21 No. 7, pp. 600-619.

Salehzadeh, R., Shahin, A., Kazemi, A. and Barzoki, A.S. (2015), "Is organizational citizenship behavior an attractive behavior for managers? A Kano model approach", *Journal of Management Development*, Vol. 34 No. 5, pp. 601-620.

Schnake, M.E., Williams, R.J. and Fredenberger, W. (2007), "Relationships between frequency of use of career management practices and employee attitudes, intention to turnover, and job search behavior", *Journal of Organizational Culture, Communications and Conflict*, Vol. 11 No. 1, p. 53.

Seibert, S.E. and Kraimer, M.L. (2001), "The five-factor model of personality and career success", *Journal of Vocational Behaviour*, Vol. 58 No. 1, pp. 1-21.

Sen, C., Mert, I.S. and Abubakar, A.M. (2021), "The nexus among perceived organizational support, organizational justice and cynicism", *International Journal of Organizational Analysis*, Vols ahead-of-print Nos ahead-of-print, doi: 10.1108/IJOA-11-2020-2488.

Setyaningrum, R.P., Pawar, A. and Pujiono, A. (2021), *The Influence of Work Motivation, Career Development, and Work Environment on Organizational Citizenship Behavior (A Case of Mold Manufacturing Unit of PT*, Bumimulia Indah Lestari, The Branch of Cikarang, Bekasi.

Shore, L.M. and Wayne, S.J. (1993), "Commitment and employee behaviour: comparison of affective commitment and continuance commitment with perceived organisational support", *Journal of Applied Psychology*, Vol. 78 No. 5, p. 774.

Smith, C.A.O.D.W.N.J.P., Organ, D.W. and Near, J.P. (1983), "Organisational citizenship behaviour: its nature and antecedents", *Journal of Applied Psychology*, Vol. 68 No. 4, p. 653.

Snape, E. and Redman, T. (2010), "HRM practices, organizational citizenship behaviour, and performance: a multi-level analysis", *Journal of Management Studies*, Vol. 47 No. 7, pp. 1219-1247.

Sullivan, S.E. and Arthur, M.B. (2006), "The evolution of the boundaryless career concept: examining physical and psychological mobility", *Journal of Vocational Behaviour*, Vol. 69 No. 1, pp. 19-29.

Tabachnick, B.G., Fidell, L.S. and Ullman, J.B. (2007), "Using multivariate statistics", Boston, MA: Pearson, Vol. 5, pp. 481-498.

Tanaka, J.S. and Huba, G.J. (1989), "A general coefficient of determination for covariance structure models under arbitrary GLS estimation", *British Journal of Mathematical and Statistical Psychology*, Vol. 42 No. 2, pp. 233-239.

Tziner, A. and Sharoni, G. (2014), "Organizational citizenship behavior, organizational justice, job stress, and work-family conflict: examination of their interrelationships with respondents from a non-Western culture", *Revista de Psicología del Trabajo y de las Organizaciones*, Vol. 30 No. 1, pp. 35-42.

Van Dyne, L., Graham, J.W. and Dienesch, R.M. (1994), "Organizational citizenship behavior: construct redefinition, measurement, and validation", *Academy of Management Journal*, Vol. 37 No. 4, pp. 765-802.

Vinkenburg, C.J. and Weber, T. (2012), "Managerial career patterns: a review of the empirical evidence", *Journal of Vocational Behavior"*, Vol. 80 No. 3, pp. 592-607.

Volmer, J. and Spurk, D. (2011), "Protean and boundaryless career attitudes: relationships with subjective and objective career success", *Zeitschrift für ArbeitsmarktForschung*, Vol. 43 No. 3, pp. 207-218.

Weng, Q. and McElroy, J.C. (2012), "Organisational career growth, affective occupational commitment, and turnover intentions", *Journal of Vocational Behavior*, Vol. 80 No. 2, pp. 256-265.

Weng, J., Lim, E.P., Jiang, J. and He, Q. (2010), "Twitterrank: finding topic-sensitive influential twitterers", *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, pp. 261-270.

Williams, L.J. and Anderson, S.E. (1991), "Job satisfaction and organisational commitment as predictors of organisational citizenship and in-role behaviours", *Journal of Management*, Vol. 7 No. 3, pp. 601-617.

Yen, H.R., Li, E.Y. and Niehoff, B.P. (2008), "Do organizational citizenship behaviors lead to information system success?: Testing the mediation effects of integration climate and project management", *Information and Management*, Vol. 45 No. 6, pp. 394-402.

Yildiz, D., Beskese, A. and Bozbura, F.T. (2015), "A hybrid decision-support model for self-managed career", *Kybernetes*, Vol. 44 No. 4, pp. 555-575, doi: 10.1108/K-12-2014-0289.

Zaleska, K.J. and de Menezes, L.M. (2007), "Human resources development practices and their association with employee attitudes: between traditional and new careers", *Human Relations*, Vol. 60 No. 7, pp. 987-1018.

Zhang, Y., Liu, X., Xu, S., Yang, L.Q. and Bednall, T.C. (2019), "Why abusive supervision impacts employee OCB and CWB: a meta-analytic review of competing mediating mechanisms", *Journal of Management*, Vol. 45 No. 6, pp. 2474-2497.

Corresponding author
Meha Joshi can be contacted at: mehajoshi83@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**

Check for updates

Thermodynamic and Comparative Analysis of Ejector Refrigeration Cycle and Absorption Refrigeration Cycle Integrated Wet Ethanol-Fueled HCCI Engine for Cogeneration of Power and Cooling

Mohd Asjad Siddiqui
Department of Mechanical Engineering,
Delhi Technological University,
Shahbad Daulatpur, Bawana Road,
Delhi 110042, India
e-mail: siddiquimohd.asjad@gmail.com

Abdul Khaliq
Department of Mechanical Engineering,
College of Engineering at Yanbu,
Taibah University,
Yanbu Al Bahr 41911, Saudi Arabia
e-mail: khaliqsb@gmail.com

Rajesh Kumar[1]
Department of Mechanical Engineering,
Delhi Technological University,
Shahbad Daulatpur, Bawana Road,
Delhi 110042, India
e-mail: dr.rajeshmits@gmail.com

This study attempted for the proposal and analysis of a combined cycle that consists of a wet ethanol-fueled and turbocharged homogeneous charge compression ignition (HCCI) engine coupled to ejector refrigeration cycle (ERC) and absorption refrigeration cycle (ARC) for the simultaneous generation of two distinct outputs, namely, power and refrigeration. Both first and second laws of thermodynamics were employed to develop a thermodynamic model, which has been applied to investigate the performance of combined cycle. Furthermore, performance of the combined cycle for ERC versus ARC was compared and assessed after altering operating parameters (turbocharger pressure ratio, turbocharger compressor efficiency, ambient temperature, and the entrainment ratio of ERC and generator temperature of ARC) to study their effect on engine power output, refrigeration load, exergy of refrigeration, and energy and exergy efficiencies of the cooling-power cogeneration cycle. Results show that the elevated pressure of turbocharger results in the enhancement of HCCI engine power and increase of the refrigeration of thermal load, simultaneously. However, the increasing ambient temperature shows the decline of HCCI engine efficiencies and energy efficiency of cogeneration, while the cogeneration cycle exergy efficiency is found increasing. Furthermore, the results are reported for the refrigeration performed by lithium bromide–water (LiBr–H_2O)-operated ARC, and R134a- and R290-operated ERC, respectively. Mapping of exergy destruction for the presented cogeneration cycle discovered HCCI engine, boiler of ERC, generator of ARC, and catalytic convertor as the components of significant exergy destruction. Entrainment ratio and type of refrigerant employed in ERC and the generator temperature of ARC show a marginal impact on the COPs of these cycles. [DOI: 10.1115/1.4051632]

Keywords: HCCI engine, energy efficiency, exergy efficiency, absorption refrigeration cycle, ejector refrigeration cycle

1 Introduction

Establishment of an environmentally benign sustainable society demands the improvement in thermal efficiency of land transportation vehicles. On average two-thirds of the primary energy supplied to internal combustion engines is lost via heat rejection to atmosphere [1–3]. This significant loss of energy to environment asking for the recovery of this rejected heat that would decrease the fuel consumption and reduce the environmental degradation caused because of the fossil fuel consumption. In view of conserving the fossil fuels and to protect the living environment, many investigators attempted the problems related to the design and development of renewable fueled engines [4,5]. Among the renewable fuels, ethanol has been found as one of the promising options as it has higher flame speed and higher octane number, and it is oxygenated fuel due to which it provides high rate of combustion efficiency and also reduced the level of the pollutant emissions and greenhouse gases. Therefore, ethanol-fueled engines can be considered as an alternative to meet the aim of achieving higher engine efficiency at lower emissions [6]. Ethanol can be produced from various agricultural products like corn and sugarcane, and studies carried out in regard of the production of ethanol from corn through energy balance reports that significant amount of energy is required to produce ethanol for water removal processes (dehydration and distillation), which consumes 37% of the energy that ethanol contains [7]. Therefore, efforts to utilize wet ethanol directly as fuel in combustion engines was considered as an integral part of research, and from some of the investigations, it is revealed that this may lead to considerable increase in net gain of energy from 21% to 55% [8,9]. Utilization of wet ethanol had its challenges in traditional modes of combustion (spark ignited (SI) and compression ignited (CI)) and hence requires a unique combustion process that combines the characteristics of both SI and CI engines [10]. In this regard, homogeneous charge compression ignition (HCCI) engines have been discovered as an option to burn low-grade fuels for land transportation purpose because these engines have the distinction of combining the advantages of both SI and CI combustion processes. HCCI engines utilize premixed and lean fuel–air

[1]Corresponding author.

Manuscript received September 2, 2020; final manuscript received June 20, 2021; published online July 29, 2021. Assoc. Editor: Andrew C. Nix.

Copyright © 2021 by ASME

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/14/4/041003/6734355/tsea_14_4_041003.pdf by Delhi Technological University user on 01 September 2021

mixture at high intake temperature, and the mixture is thermally autoignited following the higher compression ratio. This results in the drastic reduction of the most harmful emissions of NO_x and particulate matter (PM) while providing a higher thermal efficiency [11–13]. Because of the characteristics of utilizing low-grade fuels, which can be autoignited through the process of compression, HCCI engines are expected to exploit wet ethanol more effectively than traditional engines and hence can be considered as a suitable mode of combustion for the direct utilization of wet ethanol as a primary fuel. Analysis joining quality with quantity of energy by combining the first law of thermodynamics with the second law can be used to estimate the quality of energy resource, which transform into useful work. This thermodynamic modeling assist in evaluating the exergetic losses and the cycle efficiencies based on energy and exergy [14,15]. Simultaneous applications of the first and second law applied to investigate the performance of wet ethanol-fueled HCCI engines [16–19] discovered that the significant amount of exergy accompanied by engine exhaust heat is expelled, which goes unutilized and leads to inefficient and environmentally harmful engine performance. This waste heat can be considered as the potential energy source for driving the heat-operated refrigeration and air conditioning systems with a view to reduce the environmental impact and improve the system efficiency. Analysis of these systems from exergy point of views can provide a tool for their true theoretical performance assessment. Ejector refrigeration cycle (ERC) and absorption refrigeration cycle (ARC) are heat-operated nontraditional cooling systems, which can be employed for converting the engine exhaust heat into the refrigeration effect. ERCs have been in use since many decades for air conditioning of residential buildings. The schematic and details of ERC applied for cold production can be seen in Refs. [20,21]. In the field of land transportation, implementation of ejector technology for refrigerating the thermal load of air conditioner or cabin cooling is scarce, and hence, only few of the investigations are reported in this regard. Jaruwongwittaya and Chen [22] investigated the feasibility of employing a two-stage ejector refrigeration system in a bus, and their results discovered that replacement of conventional vapor compression system with the ejector cooling system can be suitably employed as an alternative. Zegenhagen and Ziegler [23] presented the assessment of an engine exhaust heat-driven jet-ejector cooling system integrated to turbocharged gasoline engines. They determined the energy efficiency and charge air temperature for the use of R134a as the refrigerant for ejector cooling. Unal [24] worked on the application of the exhaust heat-driven two-phase ejector cooling system for reducing the load and fuel consumption of the engine of passenger bus. The author states that the application of two-phase ejector as an expansion valve can enhance the performance of the bus air conditioning system. Zhang et al. [25] developed an exhaust heat recovery unit for the automobile integrating to the ejector refrigeration cycle. They analyzed the performance of the ejector cooling system and recommended a suitable range of operating parameters for the efficient operation of combined system. Another suitable option to recover the engine exhaust heat is the absorption refrigeration system in which the fluid pairs lithium bromide–water (LiBr–H_2O) and NH_3–H_2O are normally used to produce cooling through these systems. Since a significant portion of energy supplied to internal combustion engines appears as exhaust waste heat and the use of air conditioners in vehicles arrives at the status of essential commodity for modern transportation. In this context, some theoretical and experimental studies are reported in the literature regarding the use of absorption refrigeration cycle to produce cooling through the recovery of engine exhaust heat. Koehler et al. [26] developed a prototype of ARC for truck refrigeration utilizing the exhaust gas. On the basis of the description of simulated results, they reported that system performance could be improved nearly by double after employing the ARC. Manzela et al. [27] presented the experimental investigations of an ammonia–water absorption refrigeration system driven by engine exhaust gases. They evaluated the effect of exhaust gas enthalpy and the employment of ARC on engine performance and exhaust emissions and discovered that the application of ARC results in the decrease of specific fuel consumption and reduction in carbon monoxide emission.

No comparative study on the use of exhaust gas waste heat-operated cooling systems bottoming with the HCCI engine fueled by wet ethanol has been found in the literature. Therefore, it is necessary and meaningful to develop and analyze the systems that can be powered by engine exhaust gases and produce cooling for air conditioning of renewable fueled vehicles. In this study, engine operating in the HCCI mode and fueled by wet ethanol was proposed to be bottomed with the ERC and ARC, separately, to recover the exhaust heat for refrigerating the thermal load of vehicle air conditioning. A comparative thermodynamic analysis of the proposed cooling-power cogeneration was conducted by considering the quality and quantity of energy transfers during the energy conversion processes of the cycle. Energetic and exergetic performances of cogeneration cycle were assessed by altering the following parameters: turbocharger pressure ratio, turbocharger compression efficiency, and ambient temperature. In addition, the distribution of fuel energy supplied in terms of energy produced and energy lost as well as the breaking down of exergy of fuel supplied into the exergy produced, exergy destruction in the major components of cogeneration cycle, and the loss of exergy due to thermal exhaust to ambient is also computed, graphed, and discussed. The contribution of exergy destruction is scrutinized and debated with regard of cycle performance improvement. Results are derived for the employment of R134a as the refrigerant of ERC and LiBr–H_2O mixture as the working fluid pair for the ARC. Furthermore, coefficient of performance (COP) of both ERC and ARC were computed with the variation in the entrainment ratio and generator temperature, respectively.

2 System Description

In this article, investigation on wet ethanol-fueled and turbocharged HCCI engine coupled with ERC/ARC is performed. Figure 1 describes the working of HCCI engine fueled by wet ethanol and combined with the ejector refrigeration cycle. The engine is turbocharged and ethanol fired whose main specifications are reported in Table 1. Ambient air (1) enters the compressor (Comp1), which increases its pressure and temperature. The compressed air is heated (2–3) in regenerator (Reg1) and mixed with wet ethanol (4) injected into a fuel vaporizer (FV1), where it vaporizes and forms a homogeneous mixture of ethanol–water–air (5) and enters the HCCI engine. The ethanol–water–air mixture inducted into the cylinder and mixes with hot residual gases within the cylinder and compressed by the piston (1′–2′) as shown in Fig. 3. Charge gets heat through compression, and this leads to ignition and combustion (2′–3′) and generate power (W_{Power}). Then, combustion engine exhaust gases (6) enter the catalytic converter (CC1) and leave the convertor (7) at a higher temperature due to releasing of heat during combustion of carbon monoxide and unburned hydrocarbon, which were not burned in the combustion chamber. Less harmful gases from catalytic converter (CC1) at higher temperature enter the turbine (T1) and generate power (W_T) to drive the turbocharger. The exhaust gases leave the turbine (8), and after having exchanged, heat (8–9) with compressed air (2–3) in the regenerator (Reg1) is delivered to the boiler (Boi) for vaporizing (18–11) the refrigerant in the ejector refrigeration cycle. The vaporized refrigerant (11) at high pressure and temperature (primary vapor) enters the converging-diverging nozzle of ejector (Eje). The high-velocity vapor at the nozzle exit (11a) creates a high vacuum at the entry of mixing chamber, where it entrains secondary vapor from the evaporator (16) and mixes with the primary stream. In the mixing chamber, both primary (11a) and secondary vapors (16a) are mixed together, and there, it becomes a homogeneous mixture stream. On entering the constant area section, a normal shock wave occurs, which results in the increase of pressure. After the shock, the mixed stream

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/14/4/041003/6734355/tsea_14_4_041003.pdf by Delhi Technological University user on 01 September 2021

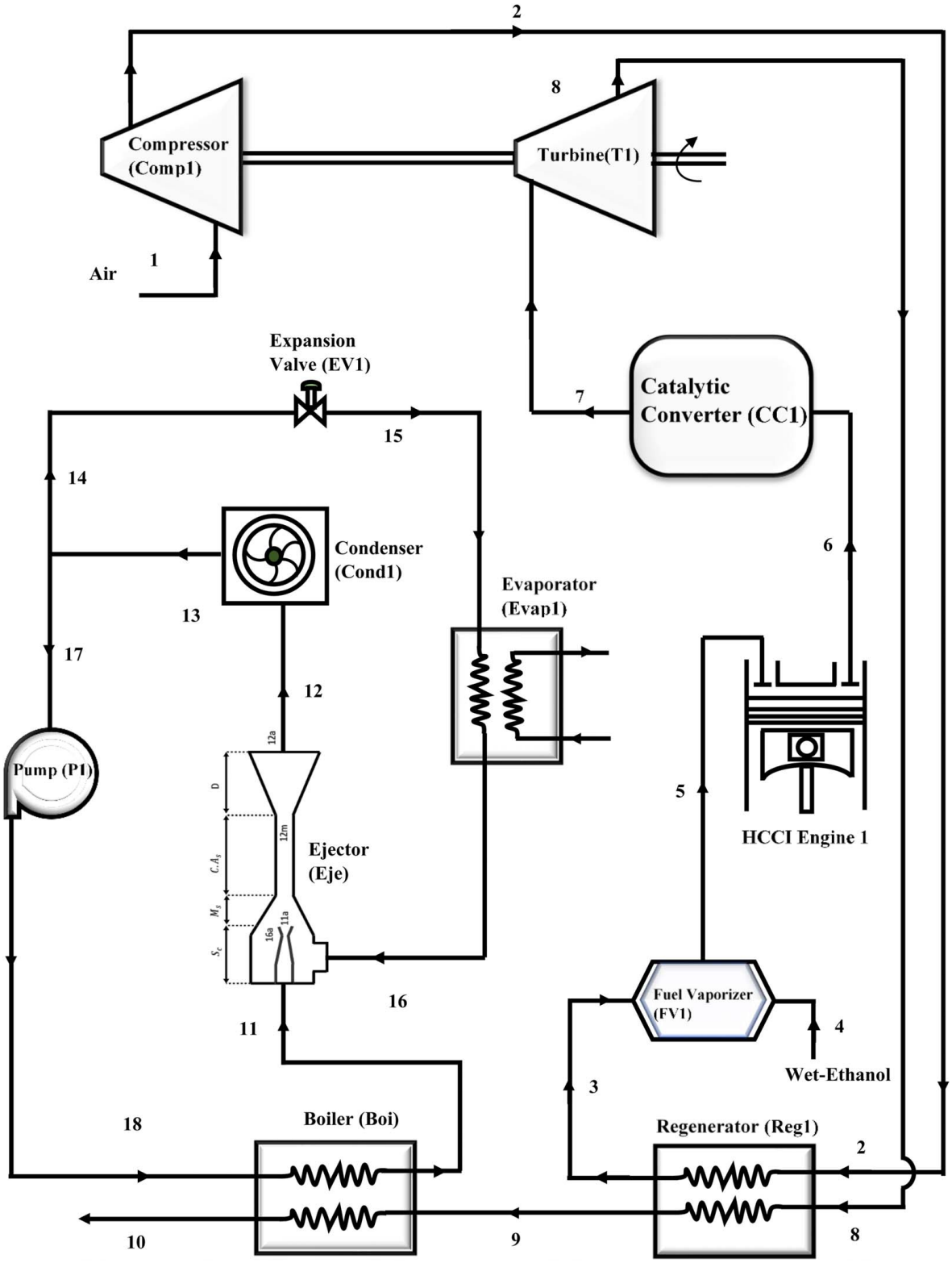


Fig. 1 System layout of wet ethanol-operated HCCI engine combined with ejector refrigeration cycle

(12 m) decelerates in the diffuser. The mixed stream (12a) from the ejector enters the condenser (Cond1) where it changes to saturated liquid refrigerant (12–13). One part of this stream as a working fluid goes to the evaporator (Evap1) to produce the desired cooling effect (15–16) through the expansion valve (EV1) (14–15), and the other part of this stream (17) is pumped back (P1) to the boiler. In Fig. 2, the turbine (T2) exhaust (8) through the regenerator delivers to the generator (Gen) (9–10) for vaporizing (25–19) the solution of LiBr–H_2O. The water vapor absorbed in the absorber (Abs) by the solution is driven off and flows to the condenser (Cond2) (19), while the LiBr–H_2O solution returns to the absorber (28) through heat exchanger (HE) (26–27) and an expansion valve (EV2b) (27–28). The vaporized refrigerant is condensed in the condenser (19–20) and then flows through an expansion valve (EV2a) (20–21) to

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/14/4/041003/6734355/tsea_14_4_041003.pdf by Delhi Technological University user on 01 September 2021

Table 1 Engine specifications [9]

Volume swept per cylinder	2400 cm^3
Cylinder diameter	13.7 cm
Piston displacement	16.5 cm
Length of connecting rod	26.2 cm
Ratio of intake to clearance volume	16:1
Speed of engine	30 rps
Engine volumetric efficiency	100%
Ratio of actual fuel–air mixture to chemically correct fuel–air mixture	<0.5
Combustion process in HCCI engine	Chemical kinetics applied to control the bulk auto ignition process

receive heat load at the evaporator (Evap2) (21–22). A LiBr–H_2O solution leaves the generator (26) and enters the absorber (28) to absorb low-pressure water vapor from the evaporator. The pump (P2) receives (23) a LiBr–H_2O solution from the absorber, increases the pressure of the strong solution (24), and pumps back to the generator (25) after passing through the HE.

3 Working Fluid Thermodynamic Properties

In the analysis and design of alternative fueled internal combustion engine, thermodynamic properties of working fluids are necessary to investigate its energetic and exergetic performances. In the ideal engine cycle, the working fluid that occurs for processes in

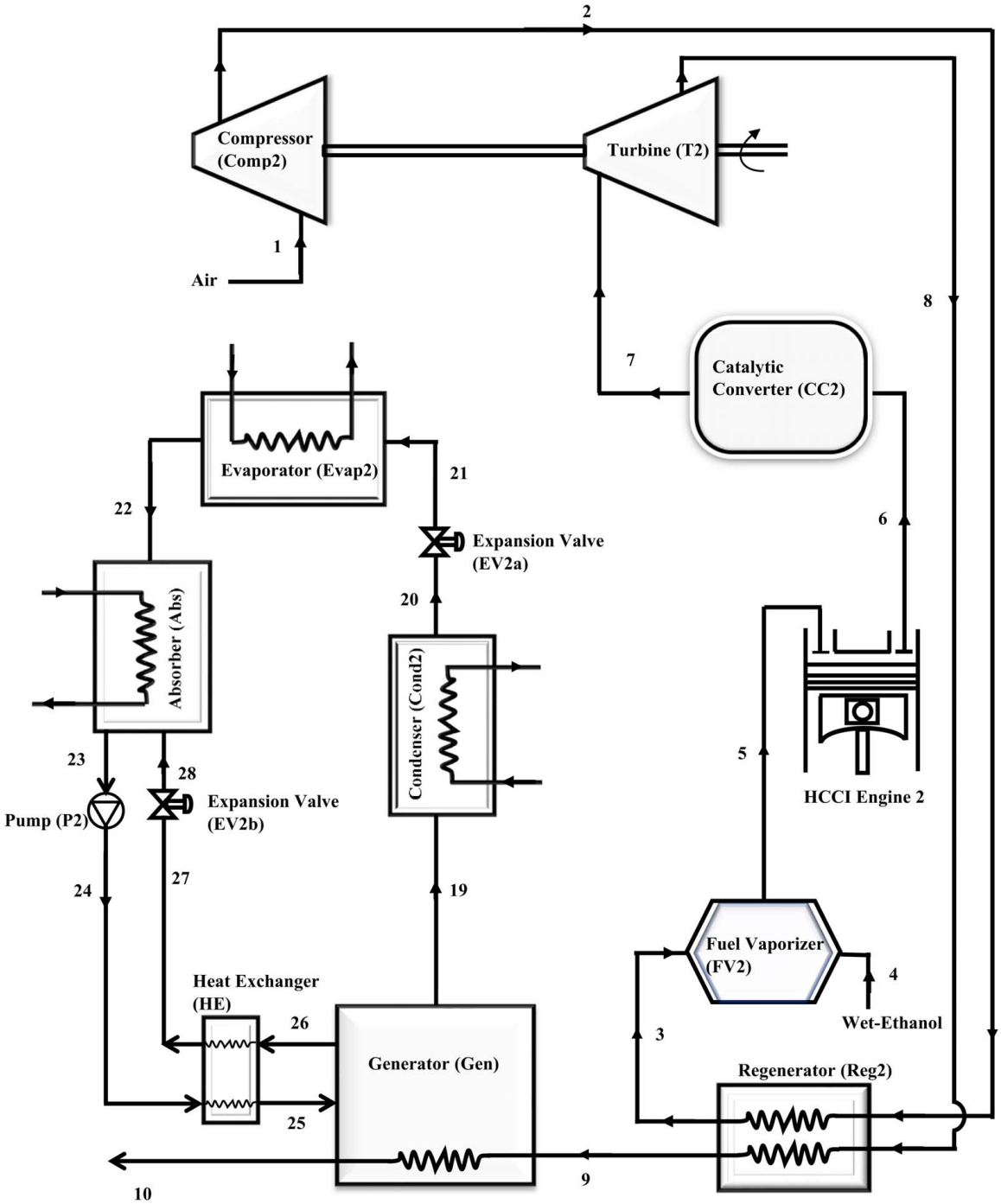


Fig. 2 System layout of wet ethanol-operated HCCI engine integrated to absorption refrigeration cycle

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/14/4/041003/6734355/tsea_14_4_041003.pdf by Delhi Technological University user on 01 September 2021

the cylinder is considered simply as air. However, in actual cycles, the working fluid consists of several gaseous species.

The specific heat of gases varies with temperature in real engine application, and its variation with pressure is negligible. In the present model of molar specific heat $\bar{C}_p(T)$ variation with temperature, the coefficients (a_1–a_5) for each species are taken from the literature in the range of temperature from 300 K to 3000 K [28].

$$\bar{C}_p(T) = \bar{R}(a_1 + a_2T + a_3T^2 + a_4T^3 + a_5T^5) \tag{1}$$

The enthalpy and entropy of the gaseous species i in terms of specific heat capacity may be expressed as follows [29]:

$$\bar{h}_{i_{(T)}} = \bar{h}_{i_{(T_0)}} + \int_{T_0}^{T} \bar{C}_{p_i}(T)dT \tag{2}$$

$$\bar{s}_{i_{(T,p)}} = \bar{s}_{i_{(T_0,p_0)}} + \int_{T_0}^{T} \frac{\bar{C}_{p_i}(T)}{T}dT - \bar{R}ln\left(\frac{y_i p}{p_0}\right) \tag{3}$$

The intake air, combustion products, and engine exhaust gases are considered to be a mixture of various species. Below are the equations that can be utilized for the computation of molar specific enthalpy, molar specific entropy, and molecular mass of the mixture [29].

$$\bar{h}_{mix_{(T)}} = \sum_{i=1}^{n} y_i\bar{h}_{i_{(T)}}; \quad \bar{s}_{mix_{(T,P)}} = \sum_{i=1}^{n} y_i\bar{s}_{i_{(T,P)}}; \quad M_{mix} = \sum_{i=1}^{n} y_iM_i \tag{4}$$

In the present analysis, fuel–air mixture is stoichiometric because ethanol-containing water obtains all the required dilution. The combustion equation for burning of wet ethanol in the HCCI engine can be considered as follows [18]:

$$C_2H_5OH + 6\ H_2O + 3(O_2 + 3.77\ N_2) = 1.8\ CO_2 + 0.1\ CO + 0.05\ C_2H_5OH + 8.85\ H_2O + 0.2\ O_2 + 11.319\ N_2 \tag{5}$$

The procedure for obtaining the balanced reaction equation of an actual reaction where combustion is incomplete is not always straightforward. Actually combustion is the result of a series of very complicated and rapid chemical reactions, and the product formed depends on many factors. When fuel is burned in the cylinder of an engine, the product of reaction varies with the temperature and the pressure in the cylinder. The appearance of some carbon monoxide and unburned oxygen and fuel is due to incomplete mixing, insufficient time for complete combustion, and other factors. The combustion of wet ethanol in the HCCI engine was taking place by assuming a chemical reaction with 90% combustion efficiency.

Unburned hydrocarbons and carbon monoxide, which formed as emissions in the HCCI engine, can be further oxidized by using Pd-loaded $SiO_2 - Al_2O_3$ as the oxidizing catalyst. Because of having the stoichiometric combustion of fuel–air mixture and its operation at higher compression ratio, the temperature of HCCI engine exhaust gases that enter the catalytic convertor is considered to be enough for oxidizing the carbon monoxide and unburned fuel. Assuming the complete conversion of unburned fuel and carbon monoxide and the application of mass and energy balances, the composition and temperature of gases at the exit of catalytic convertor can be determined:

$$\sum_{Product} n_e(\bar{h}_f^{\circ} + \Delta\bar{h})_e = \sum_{Reactant} n_i(\bar{h}_f^{\circ} + \Delta\bar{h})_i \tag{6}$$

where i refers the incoming streams entering the catalytic convertor and e refers the exiting stream of products leaving the convertor [19]. The equation for the chemical reaction of engine exhaust gases in the catalytic convertor is assumed to be as follows:

$$1.8\ CO_2 + 0.1\ CO + 0.05\ C_2H_5OH + 8.85\ H_2O + 0.2\ O_2 + 11.319\ N_2 = 2\ CO_2 + 9\ H_2O + 11.319\ N_2 \tag{7}$$

3.1 Turbocharger Compressor. The air inducted is first flows through the compressor to increase its pressure and also the temperature. For a given pressure ratio and considered polytropic efficiency of the compressor, the turbocharger compressor outlet temperature can be determined after applying the following equation:

$$T_2 = T_1\left(\frac{P_2}{P_1}\right)^{\frac{(\gamma-1)}{\gamma\eta_C}} \tag{8}$$

3.2 Regenerator. Air leaving the compressor is delivered to regenerator (an intake air to exhaust gases heat exchanger) where the intake is preheated through exhaust gases. This preheating of air is required to evaporate the wet ethanol in the fuel vaporizer. The regenerator performance is determined through its effectiveness, which is given by

$$\varepsilon_h = \frac{[(\dot{m}.c_p)_2(T_3 - T_2)]}{[(\dot{m}.c_p)_{min}(T_8 - T_2)]} \tag{9}$$

The assumption $(\dot{m}.c_p)_2 = (\dot{m}.c_p)_{min}$ is taken to estimate the temperature, T_3, the air temperature at the regenerator exit, and with this relation, it is expressed as follows:

$$T_3 = T_8\varepsilon_h + T_2(1 - \varepsilon_h) \tag{10}$$

The temperature at the regenerator exit is computed by applying the energy balance over it, which is reported in Table 2.

3.3 HCCI Engine. The premixed ethanol–water mixture leaving the mixer enters the HCCI engine. The charge enters the engine, and before compression, it is mixed with hot residual gases. Due to operation of engine at higher compression ratio, the charge is heating up to the level of ignition and combustion. But autoignition process is observed in the HCCI engine for combustion, which is controlled by chemical kinetics [9]. The efficiency of combustion process is considered to be 90%. External exhaust gas recirculation is absent in this model of engine [31]. The temperature of the mixture containing the fresh charge and residual gases at the end of engine intake process ($i - 1'$) is computed by employing the expression, and the ratio of pressure for inlet/exhaust is considered to be 1.4 [32].

$$T_{1'} = \frac{T_i(1-f)}{1 - \frac{1}{(n.r_c)\left[\frac{p_e}{p_i} + (n-1)\right]}} \tag{11}$$

where $1'$ signifies the location before the process of compression of fuel–air–residual gas mixture and $P_{1'} = P_1$.

Polytropic compression stroke ($1'$–$2'$)

$$P_{2'} = P_{1'}\left(\frac{V_{1'}}{V_{2'}}\right)^n = P_{1'}(r_c)^n \tag{12}$$

$$T_{2'} = T_{1'}(r_c)^{n-1} \tag{13}$$

Assuming the heat addition process ($2'$–$3'$) as constant volume gives

$$T_{3'} = T_{2'} + q_{in}(1-f)/c_v \tag{14}$$

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/14/4/041003/6734355/tsea_14_4_041003.pdf by Delhi Technological University user on 01 September 2021

Table 2 Component wise energy balance equations for the combined HCCI engine–ERC [18,30]

Component	Balance equations
Compressor (Comp1)	$\dot{W}_{Comp1} = \dot{m}_a(h_2 - h_1)$
Regenerator (Reg1)	$\dot{m}_a\,(h_3 - h_2) = \dot{m}_{exh}\,(h_8 - h_9)$
Fuel vaporizer (FV1)	$\dot{m}_a h_3 + \dot{m}_f h_4 = (\dot{m}_a + \dot{m}_f)h_5$
HCCI engine (Compression) process,1′–2′	$(\dot{m}_a + \dot{m}_f + \dot{m}_{rg})(h_{2'} - h_{1'}) = \dot{W}_{comp} - \dot{Q}_{surr}$
Heat addition process, 2′–3′	$(\dot{m}_a + \dot{m}_f + \dot{m}_{rg})(h_{3'} - h_{2'}) = \dot{Q}_H$
Expansion process,3′–4′	$(\dot{m}_a + \dot{m}_f + \dot{m}_{rg})(h_{3'} - h_{4'}) = \dot{W}_{exp} + \dot{Q}_{surr}$
Blowdown process, 4′–5′	$\dot{m}_{exh}\,h_{4'} = \dot{m}_{exh}\,h_{5'} + \dot{Q}_{surr}$
Catalytic convertor (CC1)	$\sum_{Product} n_e(\bar{h}_f^{\circ} + \Delta\bar{h})_e = \sum_{Reactant} n_i(\bar{h}_f^{\circ} + \Delta\bar{h})_i$
Turbine (T1)	$\dot{W}_{T1} = \dot{m}_{exh}\,(h_7 - h_8)$
Boiler (Boi)	$\dot{Q}_{boil} = \dot{m}_{11}(h_{11} - h_{18})$
Pump (P1)	$\dot{W}_{P1} = \dot{m}_{18}\,(h_{18} - h_{17})$
Ejector (Eje)	$\dot{m}_{11}h_{11} + \dot{m}_{16}h_{16} = \dot{m}_{12}h_{12}$
Condenser (Cond1)	$\dot{m}_{12}(h_{12} - h_{13}) = \dot{m}_{Cond1}(h_{Cond1,o} - h_{Cond1,i})$
Expansion valve (EV1)	$h_{14} = h_{15}$
Evaporator (Evap1)	$\dot{Q}_{Evap1} = \dot{m}_{16}\,(h_{16} - h_{15})$

$$P_{3'} = P_{2'}\left(\frac{T_{3'}}{T_{2'}}\right) \tag{15}$$

Expansion stroke (3′–4′) is polytropic, and this gives

$$P_{4'} = P_{3'}\left(\frac{1}{r_c}\right)^n \tag{16}$$

$$T_{4'} = T_{3'}\left(\frac{1}{r_c}\right)^{n-1} \tag{17}$$

Polytropic blowdown (4′–5′)

$$T_{5'} = T_{4'}\left(\frac{P_{4'}}{P_e}\right)^{\frac{(1-n)}{n}} \tag{18}$$

$$P_{5'} = P_e \tag{19}$$

The constant pressure polytropic exhaust stroke (5′–6′)

$$T_{5'} = T_e \tag{20}$$

$$P_6 = P_{5'} = P_e \tag{21}$$

$$f = \frac{1}{r_c}\left(\frac{P_e}{P_4}\right)^{\frac{1}{n}} \tag{22}$$

Residual gas fraction ($f = 0.03$) is found to be lower due to large compression ratio.

The heat lost to the ambient during the processes of exhaust and blowdown is computed by assuming 9% loss of heat during these processes [31].

The formulation for calculating the chemical exergy in HCCI engine is presented as follows:

$$e^{ch}_{HCCI} = -\Delta g + \bar{R}T_0\left\{x\mathrm{O}_2 ln\frac{P^{00}_{\mathrm{O}_2}}{P^0} + y\mathrm{H}_2\mathrm{O}ln\frac{P^{00}_{\mathrm{H}_2\mathrm{O}}}{P^0} - \sum_K \mathrm{x}_k ln\frac{P^{00}_k}{P^0}\right\} \tag{23}$$

The details of the symbols used in Eq. (23) can be found in Ref. [18].

3.4 Turbine. Exhaust from the engine is expanded in the turbine to produce power for driving the compressor.

The temperature at turbine outlet can be determined by [9]

$$T_8 = T_7\left(\frac{P_8}{P_7}\right)^{\frac{(n-1)\eta_T}{n}} \tag{24}$$

4 Evaluation Criteria and System Modeling

The proposed combined cooling-power cycle is analyzed using the engineering equation solver, which is commercially available and provides a built in mathematical and thermodynamic property functions for solving the equations describing the model developed through balances of mass, energy, and exergy [33]. Each component of the system proposed is considered as a steady-state control volume, with transfer of energy by heat, work, and mass of fluid enters and leaves. The equations derived after applying the balances of mass, energy, and exergy were employed to formulate the thermodynamic model. The thermodynamic properties of the refrigerant R134a used in ERC was computed after employing the REFPROP V 9.1, and the properties of LiBr–H_2O fluid pair for ARC were taken from Refs. [34,35].

Application of the second law of thermodynamics along with the traditional first law approach leads to the analysis that joins the quality with quantity of energy during the process of energy conversion, and it quantifies the destruction and loss of exergy at various locations of the system.

For the energetic and exergetic analyses of the cooling-power cogeneration system considered, the thermodynamic model was developed based on the following assumptions [18,25,26]:

(1) All processes proceed under the steady-state condition.
(2) Effects of kinetic and potential energies are not considered during the energy conversion.
(3) Thermodynamic equilibrium is maintained between the inlet and exit of cycle key components.
(4) The exchange of heat between the components and environment is ignored.
(5) Expansion valves of both ERC and ARC are isenthalpic and pressure reduction devices.
(6) The pressure drops in the pipes are not considered.
(7) Physical exergies of the ERC fluid and LiBr–H_2O fluid pair for ARC are only considered.

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/14/4/041003/6734355/tsea_14_4_041003.pdf by Delhi Technological University user on 01 September 2021

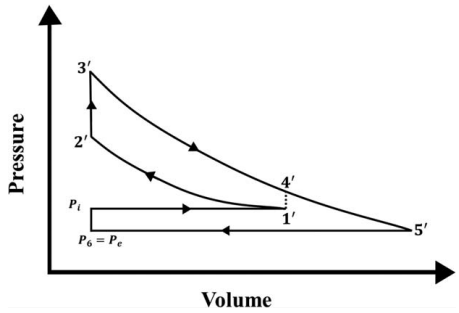


Fig. 3 P-V diagram for the HCCI engine whose thermodynamic cycle is modeled by a turbocharged Otto cycle

(8) The state of the working fluid at condenser outlet is saturated liquid.
(9) The atmospheric pressure P_0 and temperature T_0 considered for the reference environment were assumed as 25 °C and 101.325 kPa, respectively.

The traditional first law of thermodynamics along with the law of conservation of mass applied to any control volume operates at the steady state and disregards the change in kinetic and potential energies gives

$$\Sigma \dot{m}_i - \Sigma \dot{m}_o = 0 \tag{25}$$

$$\Sigma \dot{Q} - \Sigma \dot{W} + \Sigma \dot{m}_i h_i - \Sigma \dot{m}_o h_o = 0 \tag{26}$$

After applying Eqs. (12) and (13), the energy balance equations are formulated for the wet ethanol-fueled HCCI engine corresponding to its P-V diagram as shown in Fig. 3, and also for the ERC and ARC, which are employed for engine exhaust heat recovery. The formulated equations are displayed in Tables 2 and 3.

Exergy can be defined as the maximum useful work that can be obtained through a process that brings the system into equilibrium with a thermal reservoir, usually the environment.

The exergy rate balance equation that can be applied to any control volume under the steady-state operating condition can be presented as follows:

$$\Sigma\left(1 - \frac{T_0}{T}\right)\dot{Q}_j - \Sigma \dot{W} + \Sigma \dot{m}_i e_{x_i} - \Sigma \dot{m}_o e_{vx_o} - \dot{E}_{x,D} = 0 \tag{27}$$

where $\dot{m}$ is the mass flow rate of a substance. First two terms on the left-hand side of Eq. (14) indicates the exergy transfer via heat and work during a process.

$\dot{E}_{x,D}$ is the exergy destruction rate during the process. e_x is the flow exergy per unit mass of a substance, which can be broken down into the chemical and physical exergy and is given by

$$e_x = e_x^{ph} + e_x^{ch} \tag{28}$$

Physical specific exergy (e_x^{ph}) accounts for change in enthalpy of a gaseous species and loss in entropy during its transition from a specified state to the restricted dead state. It is given by the following expression:

$$e_x^{ph} = (h - h_0) - T_0(s - s_0) \tag{29}$$

The chemical exergy (e_x^{ch}) takes into account the deviations in composition of the substance from the substances occupied the reference environment. It may be expressed as [19] follows:

$$e_x^{ch} = \sum x_i(\mu_i^* - \mu_0) = -\Delta g_i^0 + \sum \bar{R} T_0 x_i ln\left(\frac{p}{p_{ref}}\right) \tag{30}$$

where μ_i is the chemical potential and g_i is the Gibbs function of species "i," and they can be evaluated at a specified temperature T and pressure p. T_0 and p_{ref} are the temperature and pressure at the reference environment, respectively. For engine applications, the reference environment temperature and pressure are taken to be as $T_0 = 298$ K and $p_{ref} = 1.013$ bar. For chemical exergy computation in engine applications, the molar composition of the environment is to be taken as follows: 20.35% O_2, 75.67% N_2, 0.03% CO_2, 3.03% H_2O, and 0.92% other substances [37].

Considering the definition of exergy under the assumptions listed, and proposed balances, the equations for the calculation of exergy destroyed during the process occurring in the main components of the proposed configuration of combined energy system are developed and presented in Tables 4 and 5.

4.1 Evaluation Criteria. Habitually, the performance of the energy conversion system is determined based on the first law efficiency, which is a measure of the percentage of the energy input that is transformed into useful energy output. The energy efficiency of the proposed cooling-power cogeneration system, which can be evaluated by applying the first law, may be presented as follows:

$$\eta_{en,comb} = \frac{\dot{W}_{BT} - \dot{W}_{Eje,\,P} + \dot{Q}_{ref}}{\dot{m}_f LHV \eta_{cc}} \tag{31}$$

where $\dot{W}_{BT}$ is the brake thermal power and $\dot{W}_{Eje,\,P}$ is the power consumed by the refrigerant pump, respectively. $\dot{m}_f$ is the rate of fuel consumption, lower heating value is the lower heating value of fuel, $\dot{Q}_{ref}$ is the rate of cooling capacity, and η_{cc} is the combustion efficiency.

The network produced in the HCCI engine, which will be available as a useful work, may be obtained after using the expression [31]

Table 3 Equations of energy balance for the components of ARC [36]

Component	Balance equations
Generator (Gen)	$\dot{m}_9 h_9 + \dot{m}_{25} h_{25} = \dot{m}_{10} h_{10} + \dot{m}_{ref} h_{19} + (\dot{m}_{sol} - \dot{m}_{ref})\, h_{26}$
Condenser (Cond2)	$\dot{m}_{ref}(h_{19} - h_{20}) = \dot{m}_{Cond2}(h_{Cond2,o} - h_{Cond2,i})$
Expansion valve (EV2a)	$h_{20} = h_{21}$
Evaporator (Evap2)	$\dot{Q}_{Evap2} = \dot{m}_{ref}(h_{22} - h_{21})$
Expansion valve (EV2b)	$h_{28} = h_{27}$
Absorber (Abs)	$(\dot{m}_{sol} - \dot{m}_{ref})h_{28} + \dot{m}_{Abs} h_{Abs,i} + \dot{m}_{ref} h_{22} + = \dot{m}_{sol} h_{23} + \dot{m}_{Abs} h_{Abs,o}$
Pump (P2)	$\dot{W}_{P2} = \dot{m}_{sol}(h_{24} - h_{23})$
HE	$\dot{m}_{sol} h_{24} + (\dot{m}_{sol} - \dot{m}_{ref})h_{26} = (\dot{m}_{sol} - \dot{m}_{ref})h_{27} + \dot{m}_{sol} h_{25}$

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/14/4/041003/6734355/tsea_14_4_041003.pdf by Delhi Technological University user on 01 September 2021

Table 4 Component wise exergy balance equations for the combined HCCI engine–ERC [18,38]

Component	Balance equations
Compressor (Comp1)	$\dot{E}_{x,D,\ Comp1} = \dot{m}_1(e_{x,1} - e_{x,2}) + \dot{W}_{Comp1}$
Regenerator (Reg1)	$\dot{E}_{x,D,\ \mathrm{Reg}1} = \dot{m}_2(e_{x,2} - e_{x,3}) + \dot{m}_8\ (e_{x,8} - e_{x,9})$
Fuel vaporizer (FV1)	$\dot{E}_{x,D,\ FV1} = \dot{m}_3 e_{x,3} + \dot{m}_4 e_{x,4} - \dot{m}_5 e_{x,5}$
HCCI engine (compression) process, 1′–2′	$\dot{E}_{x,D,\ Comp} = \dot{m}_5(e_{x,1'} - e_{x,2'}) + \dot{W}_{comp} - \dot{Q}_{surr}\left(1 - \frac{T_0}{T}\right)$
Heat addition process, 2′–3′	$\dot{E}_{x,D,\ Heat\ addition} = \dot{m}_5(e_{x,2'} - e_{x,3'}) + \dot{Q}_H\left(1 - \frac{T_0}{T_H}\right)$
Expansion process, 3′–4′	$\dot{E}_{x,D,\ \exp} = \dot{m}_5(e_{x,3'} - e_{x,4'}) - \dot{W}_{\exp} - \dot{Q}_{surr}\left(1 - \frac{T_0}{T}\right)$
Blowdown process, 4′–5′	$\dot{E}_{x,D,exh\ \&\ blowdown} = \dot{m}_5(e_{x,4'} - e_{x,5'}) - \dot{Q}_{surr}\left(1 - \frac{T_0}{T}\right)$
Catalytic convertor (CC1)	$\dot{E}_{x,D,\ CC1} = \dot{m}_6 e_{x,6} - \dot{m}_7 e_{x,7} + \dot{m}_6 e_{ch}$
Turbine (T1)	$\dot{E}_{x,D,\ T1} = \dot{m}_7(e_{x,7} - e_{x,8}) - w_{T1}$
Boiler (Boi)	$\dot{m}_{18} ex_{18} + \dot{Q}_{boi}\left(1 - \frac{T_0}{T_{ave,boi}}\right) = \dot{m}_{11} ex_{11} + \dot{m}_{boi} ex_{d,boi}$
Pump (P1)	$\dot{E}_{x,D,\ P1} = \dot{m}_{17}(e_{x,17} - e_{x,18}) + \dot{W}_{P1}$
Ejector (Eje)	$\dot{E}_{x,D,\ Eje} = \dot{m}_{11} e_{x,11} + \dot{m}_{16} e_{x,16} - \dot{m}_{12} e_{x,12}$
Condenser (Cond1)	$\dot{E}_{x,D,\ Cond1} = \dot{m}_{12} e_{x,12} + \dot{m}_{Cond1}(e_{x,Cond1,i} - e_{x,Cond1,o}) - \dot{m}_{13} e_{x,13}$
Expansion valve (EV1)	$\dot{E}_{x,D,\ EV1} = \dot{m}_{14}(e_{x,14} - e_{x,15})$
Evaporator (Evap1)	$\dot{E}_{x,D,\ Evap1} = \dot{m}_{15}(e_{x,15} - e_{x,16}) + \dot{m}_{Evap1}(e_{x,Evap1,i} - e_{x,Evap1,o})$

$$\mathrm{WD} = \left[\left(\frac{P_3' V_3' - P_4' V_4'}{n-1}\right) - \left(\frac{P_2' V_2' - P_1' V_1'}{n-1}\right)\right] \tag{32}$$

$$W_{BT} = Actual\ WD = 0.8 * WD \tag{33}$$

where W_{BT} is the actual work done, which is used to calculate brake thermal efficiency of the HCCI engine

Since the two energy outputs of the proposed cycle, power generation and rate of cooling produced, are different from the quality point of view because power is free from entropy and cooling is associated with entropy; therefore, thermal efficiency based on the traditional first law approach is not suitable for systems that produce two different outputs of energy. In this context, a relatively new term called exergy efficiency was introduced, which provides a more accurate measure of thermodynamic performance of the cycle, and it is given by

$$\eta_{ex,comb} = \frac{\dot{W}_{BT} - \dot{W}_{Eje,\ P} + \dot{E}_{ref}}{\dot{m}_f LHV \eta_{cc} \varphi} \tag{34}$$

where $\dot{E}_{ref}$ is the amount of exergy associated with the cooling capacity $\dot{Q}_{ref}$ and is given by

$$\dot{E}_{ref} = \dot{Q}_{ref}\left(\frac{T_0 - T_{Evap}}{T_{Evap}}\right) \tag{35}$$

where T_{Evap} is the evaporator temperature, T_0 is the ambient temperature, and φ is the ratio of exergy of the fuel to its energy [19].

5 Results and Discussion

The computational results after conducting the energetic and exergetic performances of the HCCI engine fueled by wet ethanol and bottoming with the ERC/ARC for cooling-power cogeneration are presented. Results are computed for the use of R134a and R290 as the refrigerants of ERC, and LiBr–H_2O as working fluid pair for ARC. In the analysis, the effects of altering the turbocharger pressure ratio, turbocharger compression ratio, ambient temperature are investigated on the efficiencies of combined cycle. The effect of entrainment ratio and generator temperature variation was also examined on the COPs of ERC and ARC, respectively.

Figure 4 represents the change in energy efficiency of HCCI engine-based combined system after altering the turbocharger pressure ratio. The HCCI engine efficiency evaluated on first law basis is found to be increasing with the increase of pressure ratio across the turbocharger. This is due to the reason higher turbocharger pressure improves the fuel–air mixing along with the increase of charge density, which increases the temperature of mixture and hence the engine power output. Furthermore, for the same turbine and compressor efficiencies and regenerator effectiveness, any increase in turbocharger pressure ratio results in the increase of engine exhaust temperature, which is an inlet temperature to the boiler, and this will increase the heat transfer to refrigerant passes through the boiler of ERC or the generator of ARC, and this in turn improves the cooling capacity of the evaporator of cycle concerned. This simultaneous increase in the output of the HCCI engine and the cooling capacity of refrigeration cycle contributes to increase in the efficiency of the cooling-power cogeneration cycle. The change of refrigerant in the ERC from R134a to R290

Table 5 Equations of exergy balance for the components of ARC cooling cycle [39]

Component	Balance equations
Generator (Gen)	$\dot{E}_{x,D,\ Gen} = \dot{m}_9(e_{x,9} - e_{x,10}) + \dot{m}_{sol} e_{x,25} - (\dot{m}_{sol} - \dot{m}_{ref}) e_{x,26} - \dot{m}_{ref} e_{x,19}$
Condenser (Cond2)	$\dot{E}_{x,D,Cond2} = \dot{m}_{ref}(e_{x,19} - e_{x,20}) + \dot{m}_{Cond2}(e_{x,Cond2,i} - e_{x,Cond2,o})$
Expansion valve (EV2a)	$\dot{E}_{x,D,\ EV,abs} = \dot{m}_{ref}(e_{x,20} - e_{x,21})$
Evaporator (Evap2)	$\dot{E}_{x,D,\ Evap2} = \dot{m}_{ref} e_{x,21} + \dot{m}_{Evap2} e_{x,Evap2,i} - \dot{m}_r e_{x,22} - \dot{m}_{Evap2} e_{x,Evap2,o}$
Expansion valve (EV2b)	$\dot{E}_{x,D,\ EV2b} = (\dot{m}_{sol} - \dot{m}_r)(e_{x,27} - e_{x,28})$
Absorber (Abs)	$\dot{E}_{x,D,\ ab} = (\dot{m}_{sol} - \dot{m}_{ref}) e_{x,28} + \dot{m}_{ref} e_{x,22} + \dot{m}_{Abs} e_{x,Abs,i} - \dot{m}_{sol} e_{x,23} - \dot{m}_{Abs} e_{x,Abs,o}$
Pump (P2)	$\dot{E}_{x,D,\ P2} = \dot{m}_{sol}(e_{x,23} - e_{x,24}) + \dot{W}_{P2}$
HE	$\dot{E}_{x,D,\ HE} = (\dot{m}_{sol} - \dot{m}_{ref})(e_{x,26} - e_{x,27}) + \dot{m}_{sol}(e_{x,24} - e_{x,25})$

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/14/4/041003/6734355/tsea_14_4_041003.pdf by Delhi Technological University user on 01 September 2021

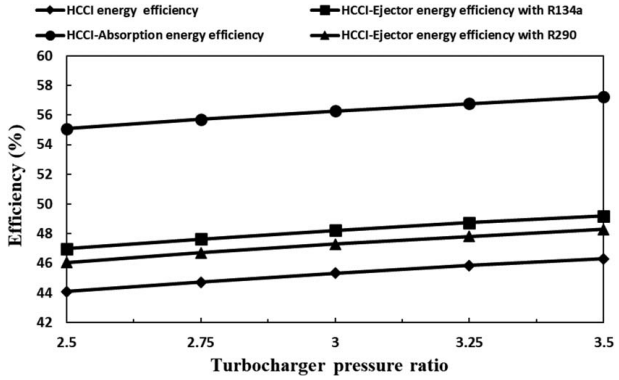


Fig. 4 Variation of energy efficiency of HCCI, HCCI-ejector system, and HCCI-absorption system with the turbocharger pressure ratio

reduces the energy efficiency of cogeneration cycle because of the considerable difference in boiling point of (−26.1 °C) for R134a and (−42.11 °C) for R290. Furthermore, lower refrigeration output for R290 (propane)-operated system attributes to the low critical temperature of propane (96.74 °C) compared to higher critical temperature of R134a (101.1 °C). Relatively a high motive flow temperature is required to produce the decent refrigeration output, and hence, the lower critical temperature of refrigerant is a hindrance in obtaining the higher output. It is further noticed that an increase in the pressure ratio across the turbocharger from 2.5 to 3.5 increases the first law efficiency of the HCCI engine from 44.09% to 46.32%, and for cogeneration, it is increased from 47% to 49.21% when R134a is used as ERC refrigerant and from 46.06% to 48.29% when R290 is the ERC refrigerant, and in case of ARC bottoming, it is increased from 55.08% to 57.25%.

Figure 5 displays the influence of turbocharger pressure ratio on the exergy efficiency of HCCI engine and the cogeneration system for both ERC and ARC, respectively. The exergy efficiency of HCCI engine is shown to be distinct than its energy efficiency, and this difference is found because of the considerable difference between the exergy and energy of fuel supplied. Since exergy efficiency of HCCI engine is the ratio of power exergy to the fuel exergy, exergy efficiency is higher than the corresponding energy efficiency of the HCCI engine. The exergy efficiency of the combined cycle (power and cooling) is found to be lower than its energy efficiency. The exergy of refrigeration is thermodynamically lower than the refrigeration capacity, and since this difference dominates over the difference between the exergy and energy of ethanol, exergy efficiency of cogeneration, which is the exergetic output to the exergy of fuel input, is appeared to be lower than the energy efficiency of cogeneration. Similar to the trend of energy efficiency of

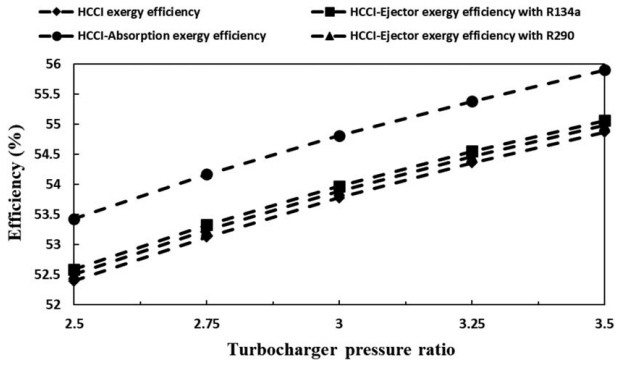


Fig. 5 Variation of exergy efficiency of HCCI, HCCI-ejector system, and HCCI-absorption system with the turbocharger pressure ratio

cogeneration, the exergy efficiency of cogeneration cycle where cooling is produced by the bottoming of ARC is found to be larger than the exergy efficiency of cogeneration bottoming with ERC.

Figure 6 displays the impact of turbocharger compressor efficiency on the energy efficiency of HCCI engine and cogeneration with both ERC and ARC. The energy efficiency of HCCI engine is found to be increased at the rise of turbocharger compressor efficiency. This is because higher turbocharger efficiency decreases the compressor power input and the turbine power, which results in the decrease of engine exhaust temperature and the higher energy efficiency of the engine. Furthermore, it is to be noted that increase in turbocharger compressor efficiency results in the decrease of energy contents of gases enters the boiler/generator of ERC/ARC, which decreases the mass flow rate of the refrigerant, and hence, the refrigeration capacity of the bottoming cycle. Since increase in the HCCI engine power dominates over the decrease of cooling capacity of ERC/ARC, the energy efficiency of cooling-power cogeneration cycle is also increased when the turbocharger compressor efficiency increases. At a given turbocharger compressor efficiency, the energy efficiency of cogeneration cycle ARC is employed as the bottoming cycle is considerably higher than ERC employed cycle for cooling. It is shown that increasing of turbocharger efficiency from 0.7 to 0.9 increases the HCCI engine efficiency from 44.51% to 45.93%, and the energy efficiency of R134a-operated ERC cogeneration increases from 47.41% to 48.83% and that of the ARC cogeneration increases from 46.48% to 47.9%, respectively.

The variation of the exergy efficiency of HCCI engine and the cogeneration cycle for ERC and ARC with the turbocharger compressor efficiency is depicted in Fig. 7. It is observed that as the turbocharger compressor efficiency increases, the exergy efficiency of the HCCI engine increases. This is because at the higher turbocharger compressor efficiency, the irreversibility decreases as in the case of compressor, and irreversibility is the difference between the actual power and the reversible power input. Since the exergy of refrigeration is significantly less than the refrigeration effect, the exergy efficiency of cogeneration cycle is little higher than the exergy efficiency of HCCI engine. It is recorded that increase in turbocharger compressor efficiency from 0.7 to 0.9 increases the exergy efficiency of the HCCI engine from 52.72% to 54.56%, and for the ARC cogeneration cycle, it is increased from 53.75% to 55.58%, respectively.

Figure 8 shows the effect of changing the ambient temperature on the energy efficiency of the HCCI engine and the cooling-power cogeneration cycle. The energy efficiency of the HCCI engine is found to be decreased with the increase of ambient temperature because of the reduction in volumetric efficiency and increase in fuel consumption in the combustion chamber at higher ambient temperature. It is further to be noted that the rate of decrease in energy efficiency of cogeneration cycle is less than the rate of decrease in the energy efficiency of the HCCI engine. This is because the refrigeration capacity of ERC/ARC starts increasing with the increase of ambient temperature due to the reduction of energy losses at heat exchanger components of ERC and ARC, but the decrease in the engine output dominates over the increase in capacity of refrigeration. Therefore, the efficiency of cogeneration diminishes gradually when the ambient temperature increases. Since the refrigeration capacity of ARC at a given temperature is greater than the refrigeration capacity of ERC, the efficiency of HCCI engine combined with the ARC is considerably greater than the case of ERC bottoming. Furthermore, it is observed that at a given ambient temperature, the exergy destroyed in the R290-operated ERC at the condenser and expansion valves is greater than the exergy destroyed in the R134a-operated ERC; therefore, the efficiency of R143a-operated ERC cogeneration is found to be higher than the efficiency of R290-operated cogeneration cycle.

The influence of ambient temperature on the exergy efficiency of HCCI engine and the cogeneration cycle was also computed and depicted in Fig. 9. The exergy efficiency of HCCI engine is

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/14/4/041003/6734355/tsea_14_4_041003.pdf by Delhi Technological University user on 01 September 2021

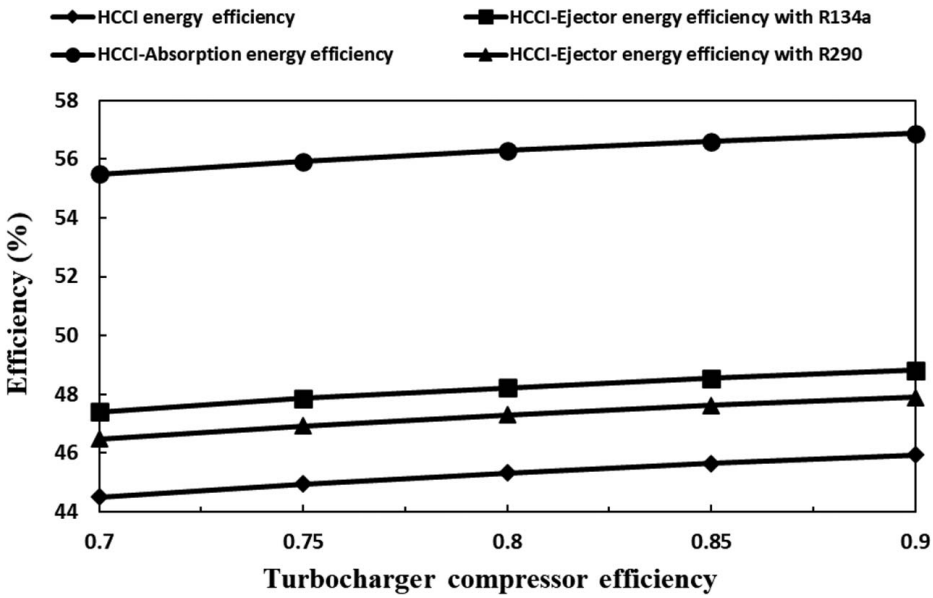


Fig. 6 Variation of energy efficiency of HCCI, HCCI-ejector system, and HCCI-absorption system with turbocharger compressor efficiency

similar to its energy efficiency but diminishes more gradually because of decrease in physical exergy of fuel at higher ambient temperature. Conversely, the exergy efficiency of the cogeneration cycle combined with ARC/ERC starts increasing because of the increase of exergy of refrigeration with the increase of the ambient temperature. By definition, exergy of refrigeration is the ratio of the refrigeration capacity to the Carnot COP between the refrigeration temperature and the ambient temperature. Since cooling capacity of ARC for the same heat input is greater than the cooling capacity of ERC, the rate of increase in the exergy efficiency of cogeneration (HCCI engine combined with ARC) is significantly higher than the rate of increase in the exergy efficiency of cogeneration (HCCI engine combined with ERC) for the same increase of ambient temperature.

The effect of change in the entrainment ratio and refrigerant type on the energetic and exergetic COPs of ERC was also examined and is shown in Fig. 10. Both energetic and exergetic COPs are found to be decreased gradually when the entrainment ratio is increased. This is because the entrainment ratio (refrigerant mass flow rate via evaporator to the flowing mass entering the ejector) increases with the increase of boiler temperature, and for the constant heat load of the boiler, this leads to the reduced mass flow of refrigerant entering the ejector and hence a lower entrained mass flow rate from the evaporator. Since the decrease in the entrained mass flow rate of refrigerant is less than the decrease of its motive mass flow rate, amount of cooling produced diminishes gradually, which shows a slight decrement in the COP of ERC. The magnitude of the exergy of refrigeration is always less than the refrigeration effect, but both depend strongly on mass of refrigerant flowing through the evaporator to generate cooling; therefore, the exergetic COP of ERC also decreases gradually with the increase of the entrainment ratio.

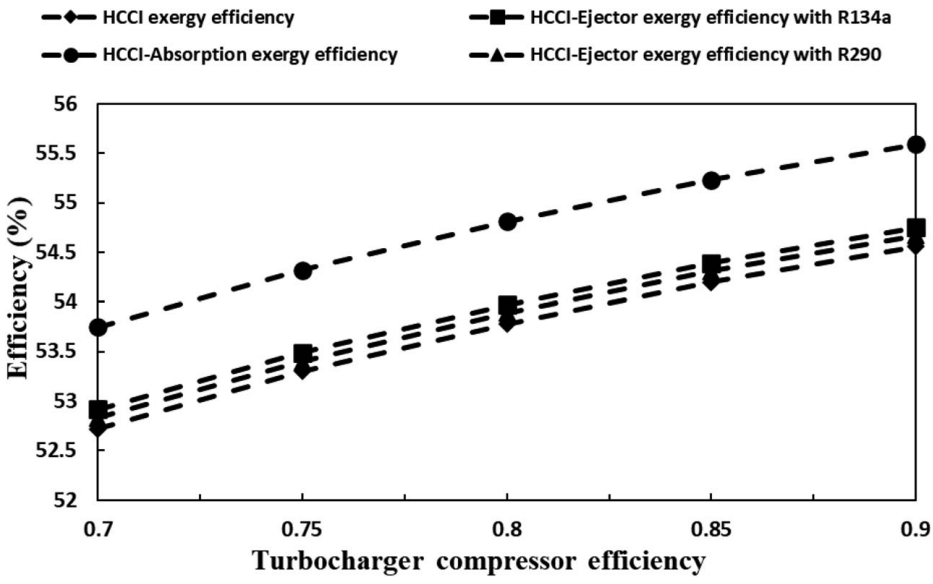


Fig. 7 Variation of exergy efficiency of HCCI, HCCI-ejector system, and HCCI-absorption system with turbocharger compressor efficiency

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/14/4/041003/6734355/tsea_14_4_041003.pdf by Delhi Technological University user on 01 September 2021

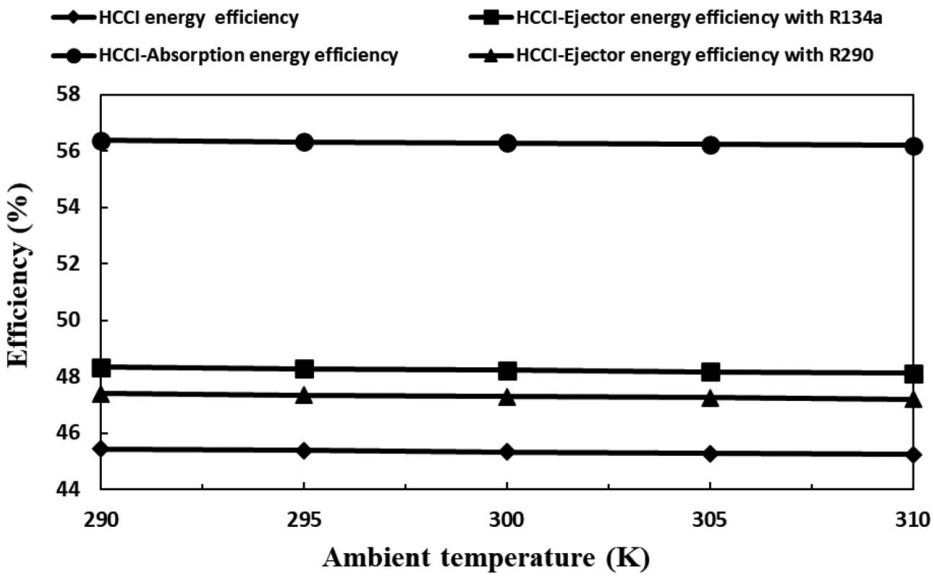


Fig. 8 Variation of energy efficiency of HCCI, HCCI-ejector system, and HCCI-absorption system with ambient temperature (K)

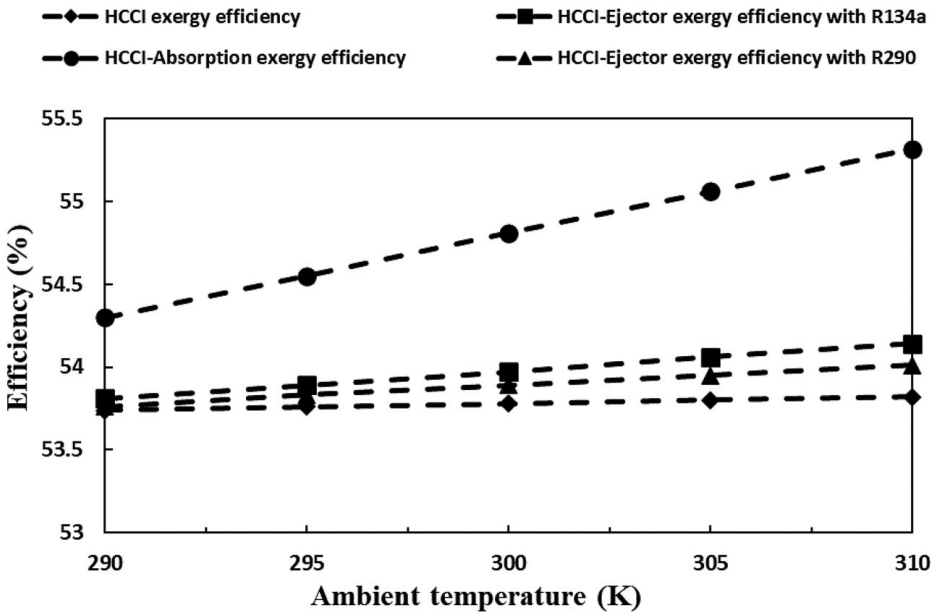


Fig. 9 Variation of exergy efficiency of HCCI, HCCI-ejector system, and HCCI-absorption system with ambient temperature (K)

Figure 11 shows the variation in COP of the ARC with the increase in generator temperature. At a constant evaporative cooling, temperature of heat transfer fluid has been increased, which results in the increase of generator temperature. This increase in temperature also increases the pressure. This in turn decreases the pressure at the outlet of expansion valve, which determines the evaporator temperature and refrigeration capacity. Therefore, the heat input to generator is increased, while keeping the cooling output as a constant in the ARC. Since the energy and exergy supplied to the generator plays an important role in the determination of COP, the energetic and exergetic COP of ARC are decreased at the increase of generator temperature. An increase in generator temperature from 359 K to 367 K declines the energetic COP of ARC from 0.7333 to 0.7299 and its exergetic COP from 0.1100 to 0.1094, respectively.

The breaking down of fuel energy supplies to combined cycle (HCCI engine–ERC) at its base line operation is illustrated in Fig. 12. From the ERC employed combined cycle where R134a is the refrigerant, it is determined that of 333.36 kW of fuel energy supplied, 130.1 kW is the produced power and 9.767 kW is the refrigeration of thermal load. About 101.34 kW of energy is lost via engine heat transfer, 45.72 kW is the energy lost via thermal exhaust to atmosphere, and 59.07 kW is the energy lost at the refrigerant condenser. In case of R290-operated ERC, because of the lower boiling point, refrigeration effect produced by R290 refrigerant is considerably less than R134a and found as 6.706 kW and energy losses at the refrigerant condenser also reduces from 59.07 kW to 56.04 kW.

The breaking down of fuel energy supplies to the combined cycle (HCCI engine–ARC) at its base line operation is illustrated in

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/14/4/041003/6734355/tsea_14_4_041003.pdf by Delhi Technological University user on 01 September 2021

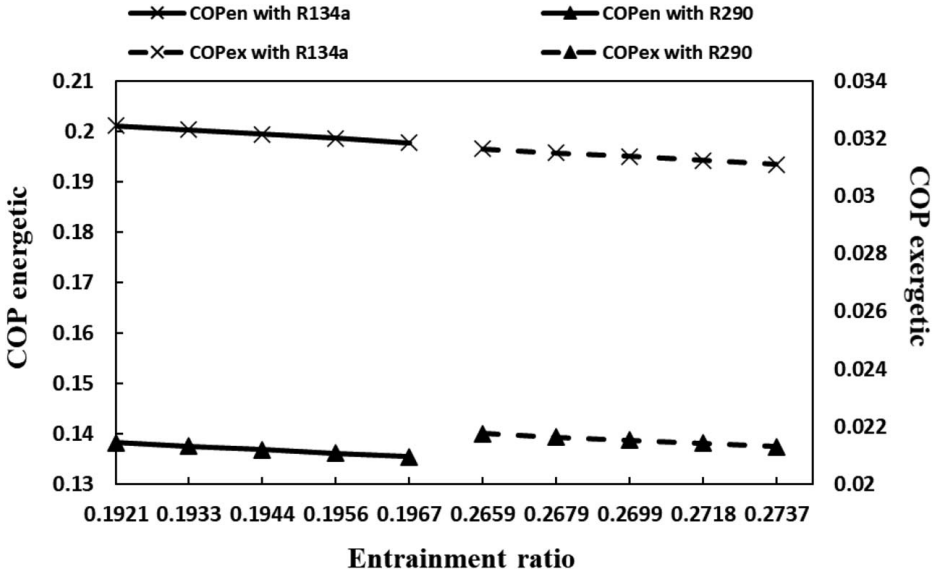


Fig. 10 Variation of energetic and exergetic COPs with entrainment ratio

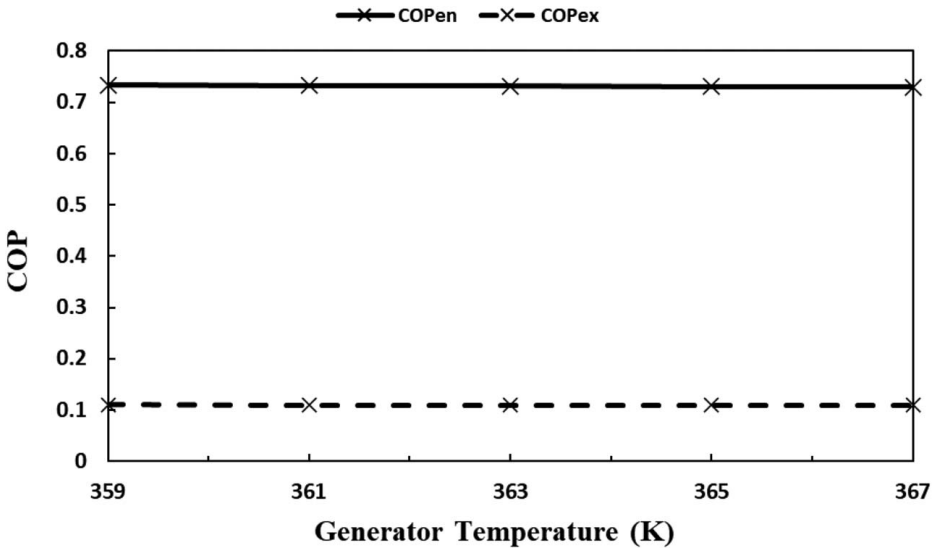


Fig. 11 Variation of energetic and exergetic COPs with generator temperature (K)

Fig. 13. The distribution of fuel energy supplied to HCCI engine-ARC shows quite different figure for the refrigeration effect and the energy losses at the refrigerant condenser. The refrigeration produced by ARC is significantly larger than ERC because of the higher boiling point of water, which act as the refrigerant in the LiBr–H_2O mixture, and it is found as 35.62 kW. The energy losses at the condenser and absorber of ARC are 38.3 kW and 46.29 kW, respectively. The distribution of fuel energy in the components of HCCI engine is unaltered by changing the type of bottoming cycle from ERC to ARC.

Employment of the principle of the decrease of exergy enunciated from the combined application of the first and second laws of thermodynamics to the developed wet ethanol-fueled combined power and cooling cycle provides a breakdown of the amount of fuel exergy supplied and is converted into engine power exergy and the refrigeration exergy, and loss of exergy as well as the exergy degraded due to generation of entropy during energy transformation in the cycle components. To this effect, the breaking down of input exergy into the exergy destroyed, loss of exergy, and the exergy produced was computed and is depicted in Fig. 14.

In case of ERC employed bottoming cycle for cold production, from the 275.5 kW of exergy supplied to combined cycle, 130.1 kW is the power exergy and 0.811 kW is the exergy of refrigeration (R134a refrigerant). Following are seen as the components of high exergy destruction: HCCI engine (61.19 kW), boiler (20.46 kW), catalytic convertor (17.89 kW), and the exergy destruction in rest of the components of the combined cycle is significantly less compared to aforementioned figures and hence not mentioned clearly in Fig. 14. The losses of exergy via engine heat transfer and thermal exhaust are found to be 37.21 kW and 7.818 kW, respectively. Replacement of R134a with R290 as the ERC refrigerant changes the exergy of refrigeration from 0.811 kW to 0.5568 kW, exergy destruction in the boiler from 20.46 kW to 21.29 kW, exergy destruction in the ejector from 3.643 kW to 2.876 kW, and exergy losses at the condenser from 1.044 kW to 1.243 kW, respectively. These differences of refrigeration exergy and exergy destruction in the boiler, ejector, and condenser arise due to the considerable difference of boiling point of two refrigerants.

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/14/4/041003/6734355/tsea_14_4_041003.pdf by Delhi Technological University user on 01 September 2021

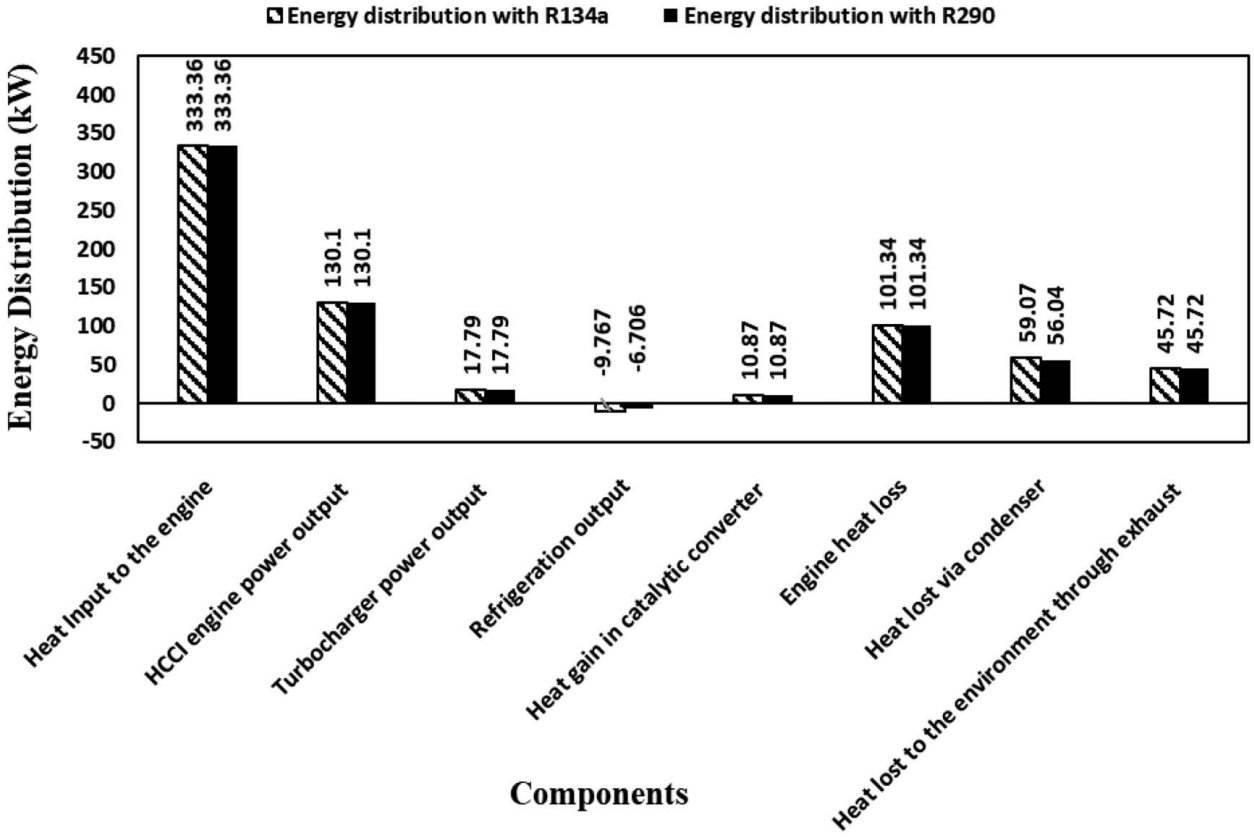


Fig. 12 Distribution of fuel energy in the HCCI engine–ERC combined cycle of cogeneration

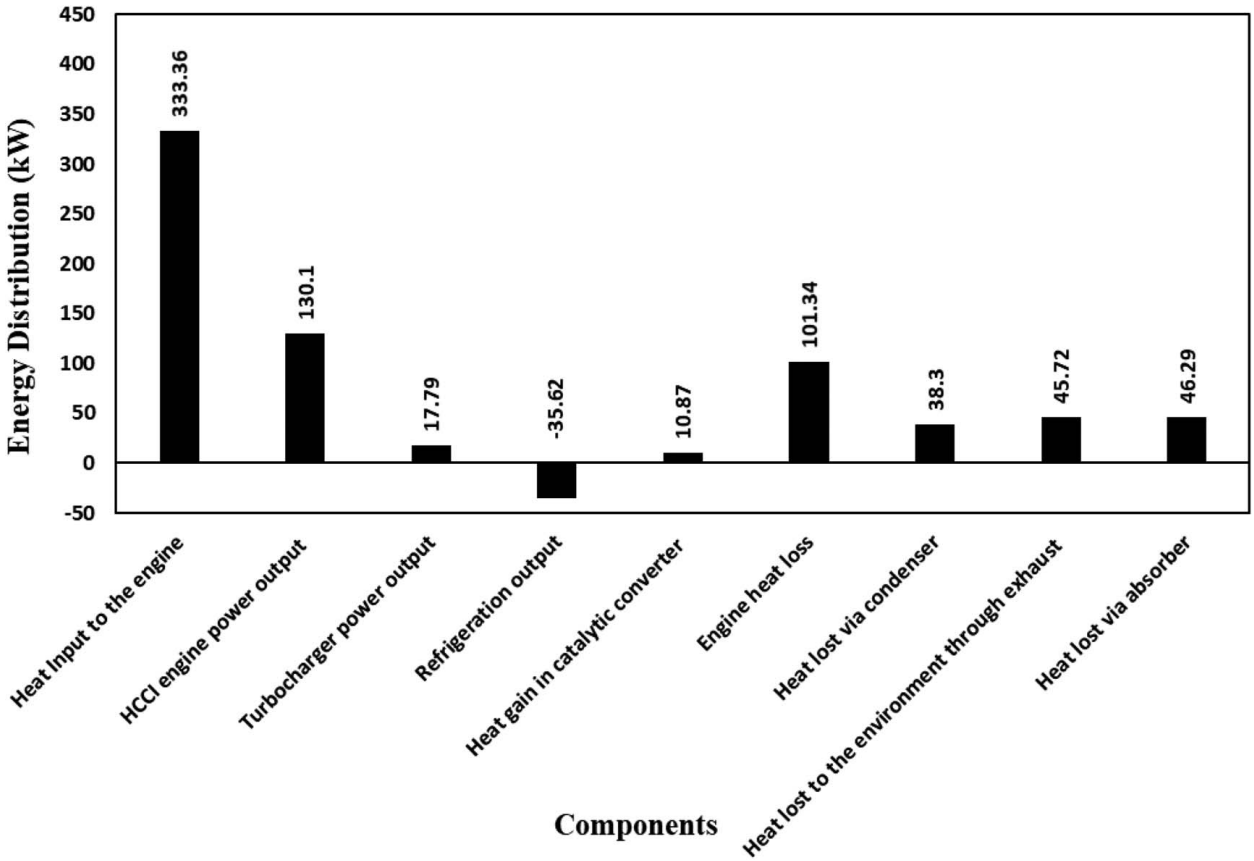


Fig. 13 Distribution of fuel energy in the HCCI engine–ARC combined cycle of cogeneration

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/14/4/041003/6734355/tsea_14_4_041003.pdf by Delhi Technological University user on 01 September 2021

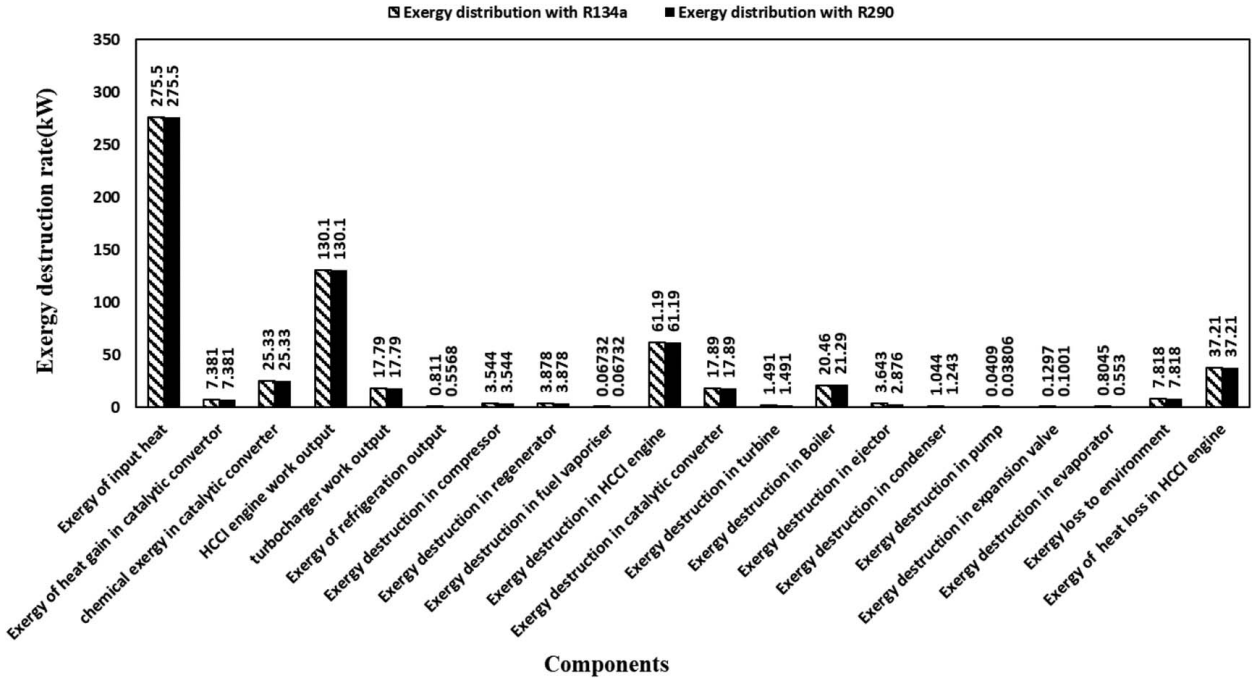


Fig. 14 Distribution of fuel exergy in the combined cycle (HCCI engine–ERC) configuration

The breaking of fuel exergy into the exergy destroyed, exergy produced, and losses of exergy via exhaust to environment for the combined cycle of HCCI engine-ARC configuration is displayed in Fig. 15. A significantly high exergy of refrigeration is found to be 2.819 kW compared to less than 1 in the case of ERC. Exergy destruction in the generator is 18.69 kW and in the condenser is 1.18 kW. These figures are different from ERC because of having a considerable difference of boiling points of R134a, R290, and water, which are utilized as refrigerants in cycles considered for the analysis. The exergy destruction in the evaporator of ARC is 0.105 kW compared to 0.805 Kw in the case of R134a ERC. This distinction arises because of the difference of the LiBr–H_2O ARC evaporator temperature above 0 °C and can be much less than 0 °C in the case of ERC evaporator. The exergy destruction in the absorber is also found to be 2.115 kW because of the mixing (source of entropy generation) of absorbent and refrigerant in the absorber. Similar to energy distribution, exergy distribution of fuel in the HCCI engine components is also found unaltered after changing the bottoming cycle from ERC to ARC.

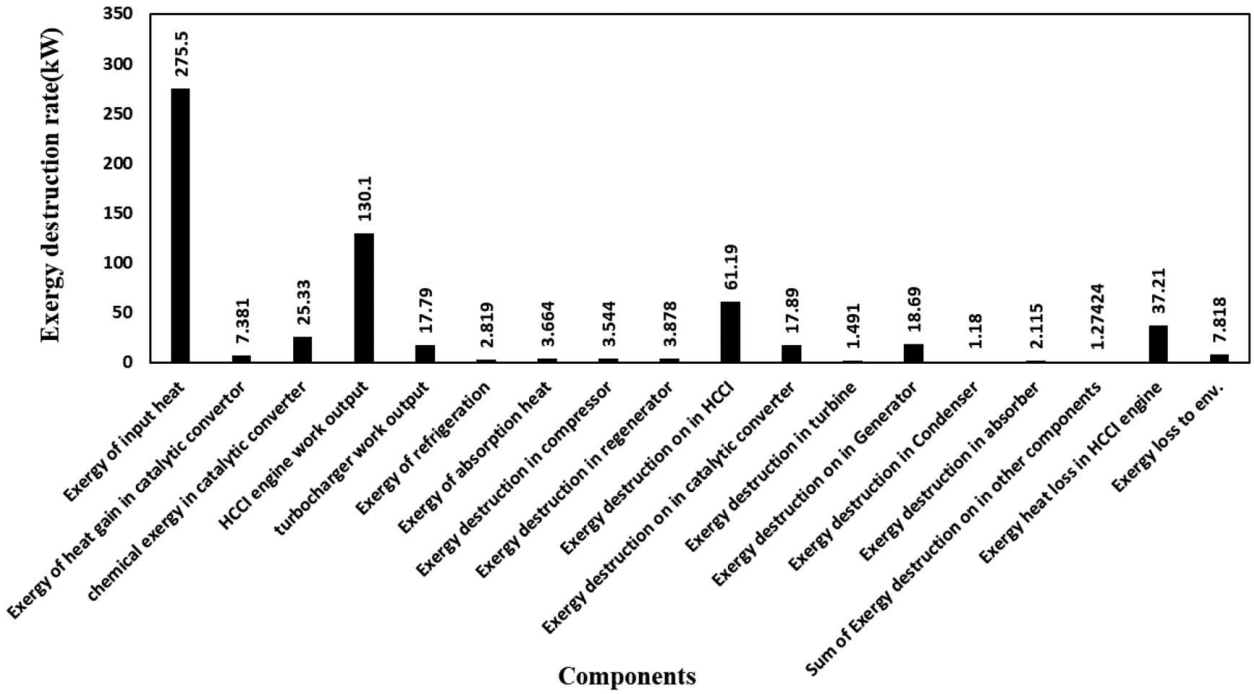


Fig. 15 Distribution of fuel exergy in the combined cycle (HCCI engine–ARC) configuration

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/14/4/041003/6734355/tsea_14_4_041003.pdf by Delhi Technological University user on 01 September 2021

6 Conclusion

This study presents a comparative study of the implementation of low-grade ethanol fuel in HCCI engine combined with two different configurations of cooling cycles for combined production of engine power and refrigeration. HCCI engine was used for power generation and the ERC/ARC powered by the engine exhaust heat were employed as bottoming cycles for the refrigeration of a thermal load. Cogeneration cycle energetic and exergetic performances were analyzed after altering the turbocharger compressor ratio, ambient temperature, compressor isentropic efficiency, entrainment ratio of ERC, and generator temperature of ARC. Breaking down of the fuel energy/exergy supplied to the combined cycle in terms of the energy/exergy produced as output, energy/exergy loss to ambient, and the exergy destruction due to irreversibilities in the cycle components is displayed. Furthermore, the comparison of computed results for the ARC ($LiBr–H_2O$) and ERC (refrigerants; R134a and R290) are presented and debated. The results revealed that operating parameters defined earlier significantly influence the cogeneration performance. The key findings obtained from this investigation can be summarized as follows:

- Both energetic and exergetic efficiencies of HCCI engine as well as of cogeneration cycle are found slightly increasing when the turbocharger compressor efficiency is increased.
- Elevation of turbocharger pressure resulting in the simultaneous increase of HCCI engine power and refrigeration of thermal load.
- The increase of ambient temperature involves the decline of energy and exergy efficiencies of HCCI engine and also the energy efficiency of cogeneration, while the exergy efficiency of cogeneration is found increasing.
- Variation of entrainment ratio and type of refrigerant employed shows a marginal decline in the energetic and exergetic COPs of ERC.
- Bottoming of cooling cycles reveals that at the system baseline operation, the refrigeration performed by ARC, R134a ERC, and R290 ERC was 35.62 kW, 9.767 kW, and 6.706 kW, respectively.
- HCCI engine (61.19 kW), boiler of ERC (20.46 kW), generator of ARC (18.69 kW), and catalytic convertor (17.89 kW) are discovered as the components of significant exergy destruction.
- Replacement of R134a with R290 as the ERC refrigerant changes the exergy of refrigeration from 0.811 kW to 0.5568 kW, exergy destroyed in the boiler from 20.46 kW to 21.29 kW, exergy destroyed in the ejector from 3.643 kW to 2.876 kW, and exergy losses at the condenser from 1.044 kW to 1.243 kW, respectively.
- The exergy destruction in the evaporator of ARC is 0.105 kW compared to 0.805 kW in the case of R134a ERC.

Findings of the current study clearly reveal the benefits of the exergy analysis over the traditional energy analysis. The comparison between the destruction of exergy at the individual component determined that HCCI engine carries a potential of major improvement in view of improving the overall efficiency of the proposed cogeneration cycle. Estimation of the fuel's energy and exergy distribution in the proposed configuration of power cooling cogeneration behaves distinctly.

Conflict of Interest

There are no conflicts of interest.

Data Availability Statement

The datasets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request. The authors attest that all data for this study are included in the paper. No data, models, or code were generated or used for this paper.

Nomenclature

g = specific Gibbs function, kJ/kg
h = specific enthalpy, kJ/kg
n = polytropic index, number of moles
p = pressure, kPa
s = entropy per unit mass, kJ/kg K
x = mass fraction
y = mole fraction
M = molecular mass, kg/kmol
R = molar gas constant, J/K mol
T = temperature, K
$\dot{m}$ = rate of flowing mass, kg/s
$\dot{Q}$ = heat transfer per unit time, kW
$\dot{W}$ = output power, kW
c_p = specific heat at constant pressure, kJ/kg
e_x = specific flow exergy, kJ/kg
e_x^{ch} = chemical exergy, kJ/kg
e_x^{ph} = physical exergy, kJ/kg
$\bar{h}_f^{\circ}$ = enthalpy of formation on molar basis, kJ/kmol
$\dot{E}x$ = exergy rate, kW

Greek Symbols

η = efficiency, %
η_{en} = energy efficiency, %
η_{ex} = exergy efficiency, %

Subscripts

a = air
Abs = absorber
Boi = boiler
BT = brake thermal
ch = chemical
$Comp1$ = compressor of ERC
$Comp2$ = compressor of ARC
$Cond1$ = condenser of ERC
$Cond2$ = condenser of ARC
$CC1$ = catalytic converter of ERC
$CC2$ = catalytic converter of ARC
D = destruction
exh = exhaust
exp = expansion
Eje = ejector
EV = expansion valve
$Evap1$ = evaporator of ERC
$Evap2$ = evaporator of ARC
f = fuel
$FV1$ = fuel vaporizer of ERC
$FV2$ = fuel vaporizer of ARC
Gen = generator
HE = heat exchanger
i = inlet
o = Outlet
$P1$ = pump of ERC
$P2$ = pump of ARC
Ph = physical
rg = residual gas
$Reg1$ = regenerator of ERC
$Reg2$ = regenerator of ARC
sol = solution
$surr$ = surrounding

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/14/4/041003/6734355/tsea_14_4_041003.pdf by Delhi Technological University user on 01 September 2021

$T1$ = turbine of ERC
$T2$ = turbine of ARC

References

[1] Armstead, J. R., and Miers, S. A., 2014, "Review of Waste Heat Recovery Mechanisms for Internal Combustion Engine," ASME J. Therm. Sci. Eng. Appl., **6**(1), p. 014001.
[2] Srinivas, K. K., Mago, P. J., and Krishnan, S. R., 2010, "Analysis of Exhaust Waste Heat Recovery From a Dual Fuel Low Temperature Combustion Engine Using an Organic Rankine Cycle," Energy, **35**(6), pp. 2387–2399.
[3] Sara, H., Chalet, D., and Cormerais, M., 2018, "Different Configurations of Exhaust Gas Heat Recovery in Internal Combustion Engine: Evaluation on Different Driving Cycles Using Numerical Simulations," ASME J. Therm. Sci. Eng. Appl., **10**(4), p. 041010.
[4] Hill, J., 2007, "Environmental Costs and Benefits of Transportation Biofuel Production From Food and Lignocelluloses-Based Energy Crops: A Review," Agron. Sustainable Dev., **27**(1), pp. 1–12.
[5] Ma, F., and Wang, Y., 2008, "Study on the Extension of Lean Operation Limit Through Hydrogen Enrichment in a Natural Gas Spark-Ignition Engine," Int. J. Hydrogen Energy, **33**(4), pp. 1416–1424.
[6] Prashant, D. K., Lata, D. B., and Joshi, P. C., 2016, "Investigation on the Effect of Ethanol Blend on the Combustion Parameters of Dual Fuel Diesel Engine," Appl. Therm. Eng., **96**, pp. 623–631.
[7] Fagundez, J. L. S., Sari, R. L., Mayer, F. D., Martins, M. E. S., and Salau, N. P. G., 2017, "Determination of Optimal Wet Ethanol Composition as a Fuel in Spark Ignited Engine," Appl. Therm. Eng., **112**, pp. 317–325.
[8] Patzek, T. W., 2004, "Thermodynamics of Corn-Ethanol Bio-Fuel Cycle," Crit. Rev. Plant Sci., **23**(6), pp. 519–567.
[9] Frias, J. M., Aceves, S. M., and Flowers, D. L., 2007, "Improving Ethanol Life Cycle Energy Efficiency by Direct Utilization of Wet-Ethanol in HCCI Engines," ASME J. Energy Resour. Technol., **129**(4), pp. 332–337.
[10] Mack, J. H., Aceves, S. M., and Dibble, R. W., 2009, "Demonstrating Direct use of Wet-Ethanol in a Homogeneous Charge Compression Ignition Engine," Energy, **34**(6), pp. 782–787.
[11] Yao, M. F., Zhang, Z., and Liu, H., 2009, "Progress and Recent Trends in Homogeneous Charge Compression Ignition (HCCI) Engines," Prog. Energy Combust. Sci., **35**(5), pp. 398–437.
[12] Christensen, M., Johansson, B., and Einwall, P., 1997, "Homogeneous Charge Compression Ignition (HCCI) Using Iso-Octane, Ethanol and Natural Gas—A Comparison With Spark Ignition Operation," SAE Technical Paper No. 972874.
[13] Mack, J. H., Dibble, R. W., Buchholz, B. A., and Flowers, D. L., 2005, "The Effect of the Di-Tertiary Butyl Peroxide (DTBP) Additive on HCCI Combustion of Fuel Blends of Ethanol and Di-Ethyl Ether," SAE Technical Paper No. 2005-01-2135.
[14] Rokopoulos, C. D., and Giakoumis, E. G., 2004, "Parametric Study of Transient Turbocharged Diesel Engine Operation From the Second law Perspective," SAE Technical Paper No. 2004-01-1679.
[15] Caton, J. A., 2012, "Exergy Destruction During the Combustion Process as Functions of Operating and Design Parameters for a Spark Ignition Engine," Int J. Energy Res., **36**(3), pp. 368–384.
[16] Boldaji, M. R., Gainey, B., and Lawler, B., 2019, "Thermally Stratified Compression Ignition Enabled by wet Ethanol With a Split Injection Strategy: A CFD Simulation Study," Appl. Energy, **235**, pp. 813–826.
[17] Saxena, S., Schneider, S., Aceves, S., and Dibble, R., 2012, "Wet Ethanol in HCCI Engines With Exhaust Heat Recovery to Improve the Energy Balance of Ethanol Fuels," Appl. Energy, **98**, pp. 448–457.
[18] Khaliq, A., Trivedi, S. K., and Dincer, I., 2011, "Investigation of a Wet Ethanol Fuelled HCCI Engine Based on First and Second Law Analysis," Appl. Therm. Eng., **31**(10), pp. 1621–1629.
[19] Khaliq, A., and Trivedi, S. K., 2012, "Second Law Assessment of a Wet-Ethanol Fueled HCCI Engine Combined With Organic Rankine Cycle," ASME J. Energy Resour. Technol., **134**(2), p. 022201.
[20] Roman, R., and Hernandez, J. I., 2011, "Performance of Ejector Cooling Systems Using Low Ecological Impact Refrigerants," Int. J. Refrig., **34**(7), pp. 1707–1716.
[21] Galindo, J., Gil, A., Dolz, V., and Ponce-Mora, A., 2020, "Numerical Optimization of an Ejector for Waste Heat Recovery Used to Cool Down the Intake Air in an Internal Combustion Engine," ASME J. Therm. Sci. Eng. Appl., **12**(5), p. 051024.
[22] Jaruwongwittaya, T., and Chen, G., 2012, "Application of Two Stage Ejector Cooling System in a Bus," Energy Procedia, **14**, pp. 187–197.
[23] Zegenhagen, M. T., and Ziegler, F., 2015, "Feasibility Analysis of an Exhaust Gas Waste Heat Driven Jet-Ejector Cooling System for Charge Air Cooling of Turbocharged Gasoline Engines," Appl. Energy, **160**, pp. 221–230.
[24] Unal, S., 2015, "Determination of the Ejector Dimensions of a Bus Air-Conditioning System Using Analytical and Numerical Methods," Appl. Therm. Eng., **90**, pp. 110–119.
[25] Zhang, H., Wang, L., Jia, L., and Wang, X., 2018, "Performance Investigation of Automobile Waste Heat Recovery System for Ejector Refrigeration Cycle," 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), May 31–June 2, Wuhan, China, pp. 400–405.
[26] Koehler, J., Tegethoff, W. J., Westphalen, D., and Sonnekalb, M., 1997, "Absorption Refrigeration System for Mobile Applications Utilizing Exhaust Gases," Heat Mass Transfer, **32**(5), pp. 333–340.
[27] Manzela, A. A., Hanriot, S. M., Cabezas-Gomez, L., and Sodre, J. R., 2010, "Using Engine Exhaust Gas as Energy Source for an Absorption Refrigeration System," Appl. Energy, **87**(4), pp. 1141–1148.
[28] Turns, S. R., and Kraige, D. R., 2007, *Properties Tables Booklet for Thermal Fluids Engineering*, Cambridge University Press, New York.
[29] Khaliq, A., Habib, M. A., and Choudhary, K. A., 2019, "Thermo-Environmental Evaluation of a Modified Combustion gas Turbine Plant," ASME J. Energy Resour. Technol., **141**(4), p. 042004.
[30] Seckin, C., 2019, "Effect of Operational Parameters on a Novel Combined Cycle of Ejector Refrigeration Cycle and Kalina Cycle," ASME J. Energy Resour. Technol., **142**(1), p. 012001.
[31] Heywood, J. B., 1998, *Internal Combustion Engines Fundamentals*, McGraw Hill, New York.
[32] Steinhilber, T., and Sattelmayer, T., 2006, "The Effect of Water Addition on HCCI Diesel Combustion," SAE Technical Paper No. 2006-01-3321.
[33] Klein, S. A., 2012, "Engineering Equation Solver (EES) for Microsoft Windows Operating Systems: Academic Professional Version: F-Chart Software," Version 9.1, Madison, WI, https://fchartsoftware.com/
[34] REFPROP, 2013, "NIST Reference Thermodynamic and Transport Properties," Version 9.1.
[35] Chua, H. T., Toh, H. K., Malek, A., Ng, K. C., and Srinivasan, K., 2000, "Improved Thermodynamic Property Fields of LiBr-H_2O Solutions," Int. J. Refrig., **23**(6), pp. 412–429.
[36] Dhahad, H. A., Hussen, H. M., Nguyen, P. T., Ghaebi, H., and Ashraf, M. A., 2020, "Thermodynamic and Thermoeconomic Analysis of Innovative Integration of Kalina and Absorption Refrigeration Cycles for Simultaneously Cooling and Power Generation," Energy Convers. Manage., **203**, p. 112241.
[37] Saxena, S., Bedoya, I. D., Shah, N., and Phadke, A., 2013, "Understanding Loss Mechanisms and Identifying Areas of Improvement for HCCI Engines Using Detailed Exergy Analysis," ASME J. Eng. Gas Turbines Power, **135**(9), p. 091505.
[38] Khaliq, A., 2017, "Energetic and Exergetic Performance Investigation of a Solar Based Integrated System for Cogeneration of Power and Cooling," Appl. Therm. Eng., **112**, pp. 1305–1316.
[39] Shokati, N., Ranjbar, F., and Yari, M. A., 2018, "Comprehensive Exergoeconomic Analysis of Absorption Power and Cooling Cogeneration Cycles Based on Kalina Part I: Simulation," Energy Convers. Manage., **158**, pp. 437–459.

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/14/4/041003/6734355/tsea_14_4_041003.pdf by Delhi Technological University user on 01 September 2021

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

Time-Borrowing Flip-Flop Architecture for Multi-Stage Timing Error Resilience in DVFS Processors

Avisekh Ghosh
Department of Electrical Engineering
Delhi Technological University,
Delhi, India
4visekh@gmail.com

Mohd Saif Naseem
Department of Electrical Engineering
Delhi Technological University,
Delhi, India
saifk5607@gmail.com

Chaudhry Indra Kumar
Department of Electrical Engineering
Delhi Technological University,
Delhi, India
indrachaudhry@dtu.ac.in

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498379

***Abstract*— Near-threshold voltage (NTV) design has been proposed for low-power VLSI designs across a wide range of applications due to its optimal Energy-Delay Product. Dynamic Voltage and Frequency Scaling (DVFS) design has acquired recent interest to increase energy efficiency during runtime. During periods of low utilization, a chip's performance and power consumption can be adjusted by slowing down the clock speed and decreasing the supply voltage, thereby bringing its operation to the NTV region. However, high dynamic variations at low voltage operation is the key design challenge and barrier for near-threshold adoption. Hence, it becomes mandatory to consider large timing margins during design to ensure high yield and robust operation. Error resilient circuits help to restore the timing margins, thereby improving performance with low power consumption. This paper presents EDTB (Error Detection with Time-Borrowing), a novel architecture for multi-stage timing error resilience which masks timing errors by borrowing time from successive pipeline stages, without slowing down the clock speed. The proposed error handling scheme doesn't require error recovery mechanisms like instruction replay or roll-back; hence it doesn't associate any error recovery penalty. EDTB-based error masking is resilient to glitches and spurious transitions due to its minimal hold time requirement. EDTB architecture shows a significant improvement of up to 34% in transistor count and 50% in clock routing over the previous architecture. The circuit schematic was prototyped in Cadence Virtuoso ADE in a 7-nm FinFET based predictive process design kit; followed by simulation, validation, and further comparisons.**

Keywords—Time-borrowing, timing error, error detection, error masking, near-threshold, DVFS, energy efficient, flip-flop, reliability, FinFET

I. INTRODUCTION

Over the last few years, VLSI technology has enriched our lives with portable, handheld, and wireless gadgets. Device scaling ensured doubling of transistors in every generation, enabling high-performance computation and network just within our palms. However, the evolution of battery technologies didn't progress that much with this trend. Therefore, the next challenge is to continue delivering high-performance throughput with reduced energy consumption. Today's handheld devices seldom perform high-performance tasks, such as playing games, watching movies, or image processing. However, most time, they are used for receiving/sending text messages, browsing the internet, social media, communication, and other less computationally expensive tasks. Therefore, the processors for the mobile devices should be capable of runtime configurability over a wide range of performance modes to support a spectrum of application scenarios. This can be achieved by incorporating the dynamic voltage and frequency scaling (DVFS) technique which can switch between a range of frequencies as well as supply voltages depending upon the current application. This ensures a reasonable operating performance with a high standby time.

The dynamic power of a CMOS circuit has a quadratic dependence over its supply voltage. Hence, as the supply voltage is decreased, the energy efficiency increases quadratically; however, the logic delay also increases leading to a decrease in operating frequency. The energy-delay product (EDP) is optimal near the threshold voltage of the transistor. The supply voltage at the minimum energy point is also known as the near-threshold voltage (NTV). In the near-threshold region, a substantial decrease in energy can be noticed without significant performance degradation. Hence, this region is favorable for applications that require high efficiency at a modest speed. However, despite the massive energy savings, the performance degradation in the NTV region can be almost half of the performance in nominal voltage. Hence the circuit must be operated in NTV only during less computation-intensive tasks.

One of the key challenges in employing near-threshold operation is the severe variations in process, voltage, and temperature (PVT) leading to timing errors, which can have a critical impact on circuit functionality and performance. A study of dynamic voltage scaling in a 65 nm CMOS process [1] shows that the variation in operating frequency over fast-slow dies could span from ±18% at nominal voltage to about ±2x at NTV. While the impact of temperature variation (0-110°C) increases the frequency spread from ±5% at nominal voltage to ±2x at NTV. It is evident that, as the supply voltage approaches threshold voltage a small variation in PVT parameters has a significant impact on logic delay or the operating frequency. Hence for NTV operation, large timing margins need to be considered during design to decrease the occurrence of timing violation and ensure reliable operation with high yield. This reduces the maximum operating frequency of the circuit and hence the expected performance from NTV operation is unjustified in a practical scenario.

Online timing error-resilient circuits help to eliminate the timing margins, thereby improving performance with low power consumption. This paper proposes a new architecture for online timing error resilience which can mask timing errors by borrowing time from successive pipeline stages.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:00:05 UTC from IEEE Xplore. Restrictions apply.

II. PRIOR WORK

Various error-resilient circuits (ERCs) have been proposed in the past which provides novel techniques and hardware architectures for online resilience to timing errors. RAZOR II [2] uses a modified latch-based flip-flop used in critical path endpoints to detect transitions in a pre-defined checking window at the current cycle of sampling failure. Replacing flip-flops with latches allows the late input signal to propagate to the output even after the positive edge thereby masking setup time violations as well as metastability issues at the triggering edge. But in contrary to flip-flops, latches have significant hold time requirements. Hence short paths need to be padded using inverter chains to eliminate hold violation. These inverter chains can be very expensive in terms of area and power. Another variant of a latch-based ERC, TDTB, and DSTB with improved power consumption was proposed in [3]. However, in this design, the duty cycle of the clock needs to be modified to offset the hold-time constraints introduced by the latch. The above techniques are successful in masking single-stage timing error, but each timing error has a corresponding recovery penalty either as roll-back or local instruction replay. This mechanism prevents the occurrence of any multi-stage timing error; however, this has an adverse effect on circuit performance since the overall throughput of the system is limited by the error rate and its associated recovery penalty [4]. Moreover, the energy gain due to NTV operation is traded off with overhead due to these error correction mechanisms. In [6] the timing error in a pipeline is masked by borrowing time from the next consecutive stage and then compensating the borrowed time by stretching the clock in the next cycle. However, apart from the requirement of a four-phase clock network, the error consolidation latency to the control block is constrained to one or half clock cycle, which is not suitable for large and high-frequency systems. In [7], each critical path halfway point is equipped with a timing violation predictor and a data arrival detection logic. The master latch clock is inverted after detecting a timing error at the mid-point of combinational logic, thereby making it transparent. But this architecture suffers from high area and power overhead with an increase in critical paths. Moreover, the half-way point selection in combinational logic becomes very complicated with increased dynamic variations in the NTV region. In ERSUT [8], the original data and its delayed copy stored in a shadow-latch are compared. An error signal is activated accordingly to select between the incorrect and delayed data by a mux. This error signal ripples through the following stages for each consecutive cycle to mask the error using a pulsed-latch, without any pipeline flushing mechanism. The method can be implemented with very low area overhead; however, glitches and spurious transitions may appear for the successive pipeline stages. The method also does not monitor for multi-stage timing error which threatens the reliability of the whole approach.

For multistage error masking, TIMBER flip-flop [4] uses a modified flip-flop with a shadow master latch and a complementary clock control circuit to mask multistage timing errors by borrowing discrete time units from the following pipeline stages. In contrast, to latch-based ERCs, which can borrow half cycle of the clock; TIMBER only borrows a small portion of time (known as recovered timing margin) from the next stage of the pipeline. Its novel double edge-sampling technique ensures glitch-free output, thereby preserving the edge triggering property of flip-flop with minimal hold time. The proposed design is an improvement over TIMBER flip-flop [4] in terms of area, power, clock overhead, and time-borrowing functionality. The comparison results are available in the VIII section.

III. MOTIVATION

Based upon the distribution of critical paths, flip-flops can be classified into three types: (i) starting at critical path, (ii) ending at critical path, (iii) starting and ending at the critical path. Flip-flops that start before the critical path have no effect of delay variation as their preceding stage is not critical. Flip-flops that end at the critical path will suffer from single-stage timing error; but it can be offset by time-borrowing from the next stage using error-masking techniques in [2, 3, 8]. Flip-flops having critical paths at both ends (starting and ending) will suffer from multi-stage timing errors, as they can't borrow time from the next stage which itself is a critical path. The single-stage timing error can arise from local dynamic variations or fast-changing global dynamic variations that can frequently occur in the NTV region. A multistage timing error may occur when successive pipeline stages become critical paths, as a result of slow-changing global variations.

This paper presents a novel architecture for online timing error resilience which can mask timing errors up to *k*-stages by borrowing time from successive pipeline stages, without any requirement of additional hardware for roll-back or instruction replay. It enables error-free operation in a DVFS system for multiple clock cycles without slowing down the clock frequency. This operation requires borrowing time from the next stage. The maximum time that can be borrowed is limited. Hence, if (*k*+1)-stage timing error occurs, a temporary slowdown of clock speed may be initiated at the system level. However, according to a study of the distribution of critical paths between flip-flops in modern processors [4], the probability of occurrence of a multistage timing error spanning more than two stages is very rare. If a design is implemented with three-stage error resilience (*k*=3), the performance loss due to this temporary slowdown of clock speed can be neglected. The design can be implemented in modern processors by replacing all the critical path flops with the proposed error detector with time-borrowing flip-flop (ETFF) and a complementary clock control circuit for every pipelined stage without a significant increase in area or transistor count.

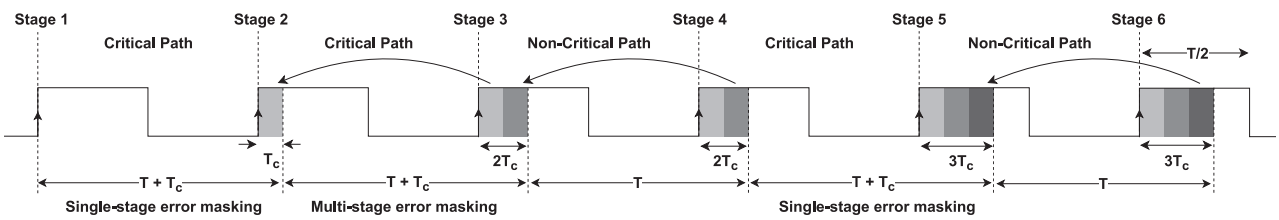


Fig. 2. . Time-borrowing flip-flops in critical and non-critical paths for a 6-stage pipeline

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:00:05 UTC from IEEE Xplore. Restrictions apply.

IV. OVERVIEW

The fundamental idea is to detect a timing error using a transition detector after the clock edge, and then mask the timing error by borrowing time from the next pipeline stage. By masking timing errors, the proposed design can recover the timing margins required to mitigate dynamic variability effects. For a single pipeline stage, the largest timing violation that could possibly occur due to dynamic variability is the timing margin that we seek to recover, and that is termed as the checking period (T_C). In other words, the checking period refers to the duration after the clock edge reserved for error detection and masking. The duration of the checking period is fixed during design and can't be changed afterward.

When the first timing error, i.e., a single-stage timing error occurs, the late-arriving data signal can cause a worst-case timing violation (T_{WC}) equal to the checking period (T_C). The proposed design can mask this single-stage timing error by borrowing a time interval equal to the duration of the checking period (T_C) from the next stage as shown in Fig. 1, stage 2. The proposed flip-flop incorporates continuous time-borrowing, i.e., it will only utilize the checking period as per the delay generated in the stage. E.g., if the delay generated is less than T_C (say $0.4T_C$), it will utilize $0.4T_C$ only and the rest $0.6T_C$ will be saved for the next consecutive stage. Continuous time-borrowing increases the hold time, but the hold time is only limited to the duration of the checking period (T_C). Now depending upon the next pipeline stage, there are two cases:

(i) If the next stage is not a critical path i.e., not affected by dynamic variability effects, then the already late-arriving data signal will not be further delayed in this stage. Hence the worst-case timing violation (T_{WC}) for this stage will be similar to the previous stage, i.e., equal to the checking period (T_C) only. This is a known as single-stage timing error. The checking period is only produced to encounter the delay incurred in the previous stage. The current stage has no contribution to T_{WC} since the path is not critical. The design need not borrow extra time in this case. No further delay due to this stage will get propagated to the next stages. This phenomenon is shown in Fig. 1, stages 4 and 6.

(ii) If the next stage is again a critical path i.e., affected by dynamic variability effects, then the late-arriving data signal can cause T_{WC} equal to twice the checking period ($2T_C$). The design will mask a two-stage timing error by borrowing an extra time interval equal to the T_C. Hence the maximum allowable delay in this stage can be as high as $2T_C$. This phenomenon is shown in Fig. 1, stage 3.

In general, the proposed design can mask timing errors up to the kth stage. However, the design can only borrow a maximum time of T/2 from the next stage (which is the length of the positive clock level). Hence, for a k-stage pipeline, the positive half cycle can be divided into k intervals, each of length T_C. The following relationship must be followed when deciding the checking period (T_C) and the maximum no. of stages (k) up to which the error will be masked.

$$T/2 = T_C \cdot k \tag{1}$$

Essentially, the hold-time of ETFF is determined by the duration of the checking period, i.e., the short paths in the design need to be padded to achieve a greater delay than the sum of the nominal hold time of the flip-flop and its checking period (T_C). But the hold time is significantly reduced as compared to latch-based error-resilient circuits [2, 3, 8], which have a hold time of half the clock cycle, i.e., T/2.

V. PROPOSED ARCHITECTURE

Fig. 2 shows the implementation of error detector with time-borrowing (EDTB) architecture in k^{th} stage for 3-stage timing error resilience. The EDTB architecture consists of two parts: (i) Error detector with time-borrowing flip-flop (ETFF), and (ii) Clock control circuit.

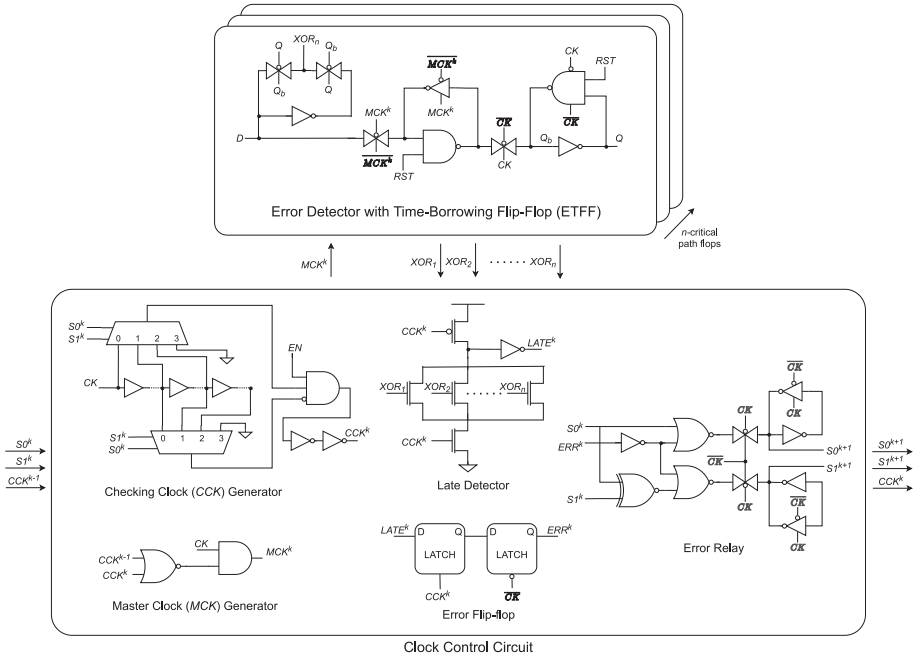


Fig. 2. EDTB architecture for k^{th} stage

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:00:05 UTC from IEEE Xplore. Restrictions apply.

A. Error Detector with Time-borrowing Flip-Flop (ETFF)

ETFF is a modified transmission gate based positive edge-triggered master-slave flip-flop. ETFFs are inserted in the place of conventional flip-flops for each critical path end-points. There are separate clocks for master and slave, named *MCK* (master clock) and *CK* (clock) respectively. *CK* is the global clock for both conventional flip-flops in non-critical paths and ETFFs on critical paths. ETFF also comprises a transmission gate based xor-gate which performs as an inequality checker. The *D* and *Q* pins of master-slave flip-flop (MSFF) are the two input signals of the xor-gate. During the positive clock cycle, the master is opaque while the slave is transparent. Hence, if *D* and *Q* pins are different, then the xor-gate will output logic-1, otherwise logic-0. The output of the xor-gate is labeled as the *XOR* signal. For *n*-critical path flops, *n-XOR* signals are generated, which are routed to the clock control circuit.

B. Clock Control Circuit

We define the clock control as those complementary circuits (other than critical path flops) which generate the checking period, handles error and relay signals for the next stage, and also enable error masking functionality in ETFF. The clock control circuit needs to be inserted for each pipeline stage. The clock control consists of: (i) Checking clock generator, (ii) Master clock generator, (iii) Late detector (iv) Error flip-flop, and (v) Error relay. The next portion of this section discusses various components of the clock control circuit.

1) Checking Clock (CCK) Generator

The checking period is a short pulse of length T_C after the positive edge of the clock as shown in Fig. 3. The checking period is generated by the checking clock (*CCK*) generator. The checking period intervals are encoded using the select input lines, S_0 and S_1. This pulse T_C can be shifted across the positive half cycle depending upon the select lines S_0 and S_1 of the muxes in the *CCK* generator. The S_0S_1 signals for the kth stage are generated from the error relay circuit of the $(k\text{-}1)^{th}$ stage. For the first stage in the pipeline (k=1), $S_0S_1^{(1)}$ is grounded. On system reset, S_0S_1 is set to *00* for all stages. The relative position of the checking clock (*CCK*) in the positive half cycle is decided by a programmable delay pulse generator. The delay is generated by a series of even no. of cascaded inverters working as a chain of buffers. The no. of buffer chain ultimately decides the no. of stages up to which the overall architecture can sustain multi-stage error. In this case, the architecture supports upto 3-stage error masking. Finally, the *CCK* is further delayed by δ using a short buffer chain consisting of 2-4 inverters. Due to this delay, the checking period ignores any signal change during this short delay period. This is important, otherwise, the system will always flag a false positive *XOR* signal due to the *Clock-to-Q* delay encountered during every *D-to-Q* transition.

Time-borrowing functionality in EDTB architecture can be deactivated by switching *EN* to zero. This gates the *CCK*, disables the clock control, thereby allowing runtime configuration of ETFF as a conventional MSFF. For time-borrowing functionality, *EN* should be tied to V_{DD}.

2) Late Detector

The late detector is a dynamic or-gate triggered by *CCK* as the clock. The *XOR* signals from *n*-ETFFs are or-ed to produce the *LATE* signal. If any of the ETFFs recognizes a transition in its input during the checking period (*CCK*=1), the *LATE* signal goes high. When the *CCK* goes low, the checking period is exhausted and the dynamic or-gate pre-charges, thereby switching the *LATE* signal to logic-0.

3) Master Clock (MCK) Generator

MCK generator is a static CMOS circuit that generates the clock signal for the master latch in ETFF. The master clock (*MCK*) is a modified version of the global clock (*CK*). For the 1st stage (k=1), input CCK^0 is grounded; hence $MCK^1 = \overline{CCK^1} \cdot CK$. Fig. 3 shows the waveform of *MCK* during the checking period T_C. When CCK^1=1, MCK^1 is gated to logic-0 and the master latch in ETFF becomes transparent. This allows the *D* signal to propagate to *Q* even when the global clock (*CK*) is high, thereby masking the timing violation. However, after the checking period is exhausted, i.e., CCK^1=0; MCK^1 simply follows *CK* and the master latch is opaque again. The ETFF behaves as a soft-edge sampling flip-flop with minimal hold time requirement, i.e., T_C. This describes the error masking property for single-stage timing error.

4) Error Flip-Flop (ERFF)

The ERFF uses a conventional MSFF with a modified master latch clock, i.e., *CCK*; while the slave latch is triggered by $\overline{CK}$. This configuration allows the ERFF to successfully capture the *LATE* signal during the checking period and update it as *ERR* signal at the falling edge of *CK* as shown in Fig. 3. The *ERR* signal is used by the error relay to generate select lines for the next stage for multi-stage error masking.

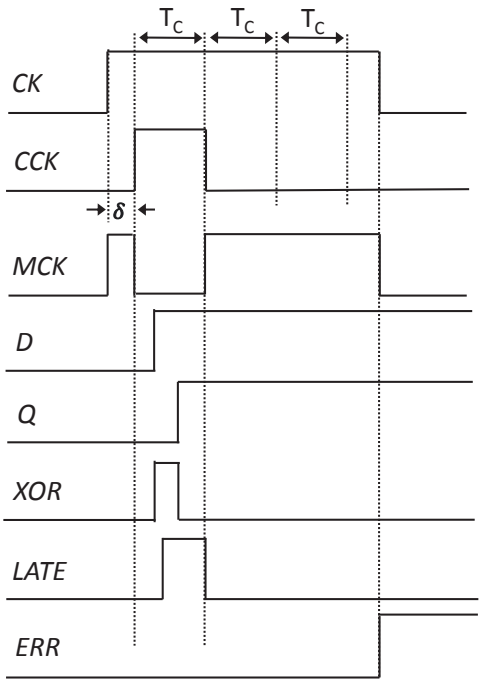


Fig. 3. Timing diagram of EDTB based single-stage error masking

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:00:05 UTC from IEEE Xplore. Restrictions apply.

5) Error Relay

The error relay circuit is inspired by the TIMBER architecture [4]. For the very first stage (k=1), $S_0S_1^{(1)}$ is set *00*. If no error occurs at stage-1, the select lines for stage-2, $S_0S_1^{(2)}$ remain *00*. If an error occurs at stage-1, then the select lines for stage-1 are incremented by one to obtain the select lines for stage-2, i.e., $S_0S_1^{(2)}$. This shifts the CCK^2 pulse from T_C to $2T_C$, thereby ensuring that the ETFFs in stage-2 can borrow an extra time interval of T_C in case a multistage timing error occurs.

VI. MULTI-STAGE ERROR MASKING

In the previous section, it was discussed how the EDTB architecture can mask single-stage timing errors. However, several architectures have been proposed and tested in the literature which can mask single-stage timing errors with much lesser area footprint and longer checking period. But these designs are not effective in masking multi-stage timing errors. The EDTB architecture's flagship feature is that it is scalable to further pipeline stages to mask multi-stage timing errors. This feature enables the processor to continue performing in NTV at the desired frequency even without any error recovery mechanisms like roll-back or instruction reply.

Fig. 4 shows the timing diagram of *CCK* and *MCK* for single-stage and multi-stage timing error cases. If stage-1 reports an error (ERR^1 = 1), $S_0S_1^{(2)}$ is incremented by error relay and the *CCK* generator mux selects the second input which shifts the checking period for stage-2, i.e., CCK^2 by one T_C interval. The effective checking period for stage-2 is now from T_C to $2T_C$, as shown in Fig. 4(b). Hence, stage-2 ignores any data change from *0* to T_C and only checks for error in the interval T_C to $2T_C$, as the former interval is the consequence of stage-1 being the critical path. The delay due to stage-2 being a critical path will only be accounted for when CCK^2 is high. There may be cases, in which stage-2 may be a short path and the actual data is captured well before T_C, or even before the positive edge. However, by shifting the checking period up to $2T_C$, we assure the design will detect the timing error even if both of the stages are critical paths.

Consequently, the shifted *CCK* also provides an extra timing margin by one T_C interval in MCK^2 to mask multi-stage timing error incurred in stage-2, as evident from comparing the waveforms of MCK^2 in Fig. 4(a) and Fig. 4(b). The Boolean equation of MCK^2 for stage-2 is given by, $MCK^2 = \overline{CCK^1 + CCK^2} \cdot CK$. In error-free case ($ERR^1$ = 0), $CCK^2 = CCK^1$ and the MCK^2 signal reduces to $MCK^2 = \overline{CCK^1} \cdot CK$, which is same as in the stage-1. However, when ERR^1 = 1, $CCK^2(t) = CCK^1(t - T_C)$ and MCK^2 becomes transparent from 0 to $2T_C$.

Note that the *ERR* signal generated from the error relay is different from the central error signal flagged to the DVFS control and error management unit. All the single-stage timing errors will be masked by borrowing one T_C interval from the next stage, but the error signal will not be flagged to the DVFS control unit. Two-stage timing errors will be masked by borrowing two T_C intervals. Similarly, k-stage timing error will be masked by borrowing k-T_C

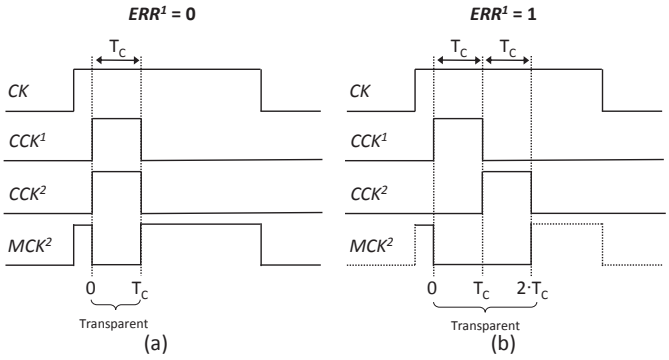


Fig. 4. . Timing diagram of CCK and MCK for (a) single-stage error masking (b) multi-stage error masking

intervals. Hence, the error signal to the DVFS control unit can be deferred until —depending upon the no. of stages an EDTB architecture is capable of masking error and the error consolidation latency incurred from the or-tree of ERR^k signals. For k-stage timing error resilience, the global error signal could be flagged at $(k\text{-}1)^{th}$ stage. As the *ERR* signal is latched at the falling edge of *CK*, the or-tree and the DVFS control unit will get a maximum of 1.5 clock cycles to take necessary action (reduce the clock frequency or add stall cycles) to mitigate the timing error rate.

The error masking techniques in [2, 3] may seem to potentially deliver a higher operating frequency due to their longer checking period. However, for each clock cycle resulting in an error, a recovery penalty is associated either as roll-back or instruction replay. With the increase in clock frequency, the error rate also increases. Hence, the clock speedup is limited by the combined penalty incurred with the error recovery mechanism following each detected error. Architectures like EDTB and TIMBER [4] can eliminate this recovery penalty even without slowing down the clock speed due to their multi-stage error masking functionality, thereby ensuring much higher system throughput in practical.

VII. SIMULATION RESULTS

The schematic of the design was prepared in Cadence® Virtuoso® Analog Design Environment with FinFET based 7 nm process design kit ASAP7 [5]. The nominal supply voltage is 0.7 V, while the near-threshold voltage (minimum EDP point) was determined to be 384 mV. Low power and low threshold MOSFET models (nmos-lvt and pmos-lvt) were used for designing with the W/L ratio being (26/20) nm/nm. The design was simulated using the Cadence® Spectre® simulator at near-threshold voltage (384 mV) and 1 GHz clock frequency. For a chain of inverters operated at the near-threshold region, almost five times improvement in the energy efficiency was observed with degradation in performance by half when compared to operation in nominal supply voltage.

Fig. 5 showcases the simulation results for single-stage as well as multi-stage error masking capability of EDTB architecture for a two-stage sequential circuit. The ETFFs in stage-1 and stage-2 are denoted by f^1 and f^2 respectively. The data, output, and error signals of the two stages are represented as D^1-D^2, Q^1-Q^2, and ERR^1-ERR^2 respectively. When the first timing error occurs at the flop f^1, it is masked

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:00:05 UTC from IEEE Xplore. Restrictions apply.

by borrowing one T_C interval at stage-1. Consecutively, the ERR^1 signal goes high at the falling edge of *CK*, thereby configuring the select lines of the CCK^2 generator for stage-2 to *01*. The two-stage timing error, occurring at the flop f^2, is masked by borrowing an extra T_c interval at stage-2. The ERR^2 signal goes high subsequently. The schematic of EDTB architecture used for simulation is the same as in Fig. 2, i.e., it is capable of masking multi-stage timing error up to three stages. Hence, the final error signal to the DVFS control unit can be flagged on the first multistage timing error, i.e., when $ERR^2 = 1$. From Fig. 5, it is evident that D^1 follows Q^1 and D^2 follows Q^2, even if *D* changes after the positive edge of the clock. This proves the time-borrowing functionality for masking single-stage as well as multi-stage timing errors. For error-free cases, the edge triggering property of the flip-flop is retained if the necessary hold conditions are met ($T_{hold} = T_C$). Since the hold time constraint is minimal (as compared to latch-based error-resilient circuits [2, 3]), the design is also resistant to spurious transitions and glitches after the checking period is exhausted.

VIII. COMPARATIVE ANALYSIS

The proposed design is compared with a functionally similar design, TIMBER [4] for a typical die operating in the NTV region. Table 1 compares the transistor count and the average power consumption, separately for flops in the critical path as well as complementary clock control circuitry. The comparison between the two is performed for three-stage error resilience with the same time-borrowing interval. Same input stimuli (*D* and *CK*) were applied to each design and their average power per stage was calculated over an operation of six clock cycles, averaged for the two stages.

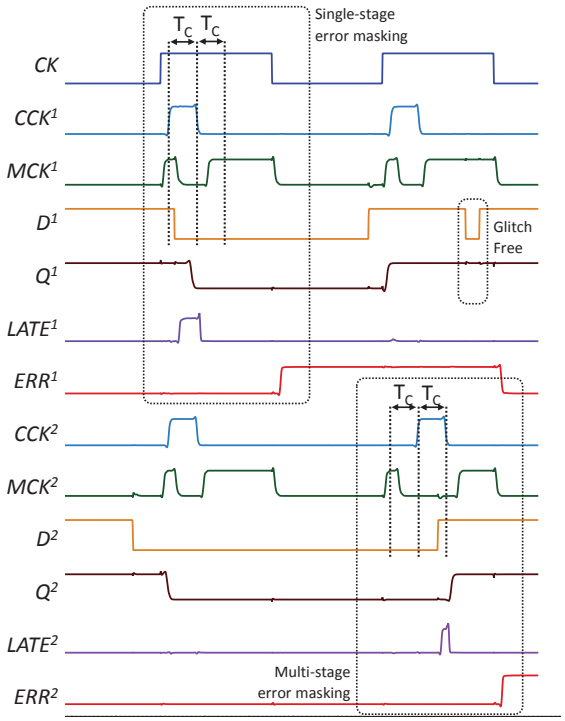


Fig. 5. Simulation waveforms for a two-stage sequential circuit with single-stage and multi-stage timing error

TABLE I. COMPARISON BETWEEN MULTI-STAGE TIMING ERROR RESILIENT CIRCUITS

Architecture	Transistor count		Avg. power (nW)	
	Flip-flop	Clock control	Flip-flop	Clock control
EDTB	30	215	219	2827
TIMBER FF	46	182	268	2750
	34.78% less	18.13% more	18.28% less	2.80% more

TIMBER uses a shadow master latch for double sampling the input data. This has an adverse effect on the area, as extra no. of transistors will be required for the shadow latch. However, ETFFs in EDTB architecture don't introduce extra circuitry to the conventional master-slave flip-flop, apart from the xor-gate comparator logic. Only the master clock is modified to mask timing errors. Hence, the proposed design enjoys a substantial improvement in area and power for each critical path flip-flops.

The clock routing is two (*CK* and *MCK*), for each critical path flops in the EDTB architecture. Whereas, four different clocks (*CK*, *DCK*, P_0, and P_1) are used in TIMBER flip-flop architecture [4]. Hence, the clock routing is cut by half in the proposed design. Apart from routing congestion, multiple clocks in design are more prone to skew, process variations, and crosstalk issues. The clock tree buffers will also account for high dynamic power consumption in TIMBER architecture.

The two designs also vary with respect to their distinct time borrowing techniques. The proposed design makes the master latch transparent during the checking period. This enables continuous time-borrowing, allowing to only borrow the minimum required time to mask the error. In the contrary, TIMBER flip-flop can only borrow discrete time units which may result in unnecessary extra time being borrowed, thereby increasing the *D-to-Q* delay for the timing error case. In [4], a dedicated TIMBER flip-flop architecture and TIMBER latch has also been proposed with continuous time-borrowing functionality, but a significant hold time is introduced as the flop/latch is transparent after the checking period; thereby accounting for glitches and spurious transition in the datapath. In EDTB architecture this is not the case as the flip-flop becomes opaque immediately after the checking period is exhausted.

The comparison shows the clock control logic in the proposed design requires more transistors and consumes more power. But the clock control circuit only exists for each pipelined stage, while flip-flops must be included in each possible critical path. The Fig. 6(a) and Fig. 6(b) shows a comparison between the two design for total transistor count and average power consumption respectively, per pipelined stage. As the no. of critical paths increases in the design, the total area, and power consumption is dominated by the flip-flops in critical paths. This can be confirmed from Fig. 6(c), where the percentage improvement in EDTB over TIMBER flip-flop architecture for both the parameters is shown. Hence, the proposed design shows a significant improvement in transistor count, clock routing as well as a remarkable enhancement in average power consumption.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:00:05 UTC from IEEE Xplore. Restrictions apply.

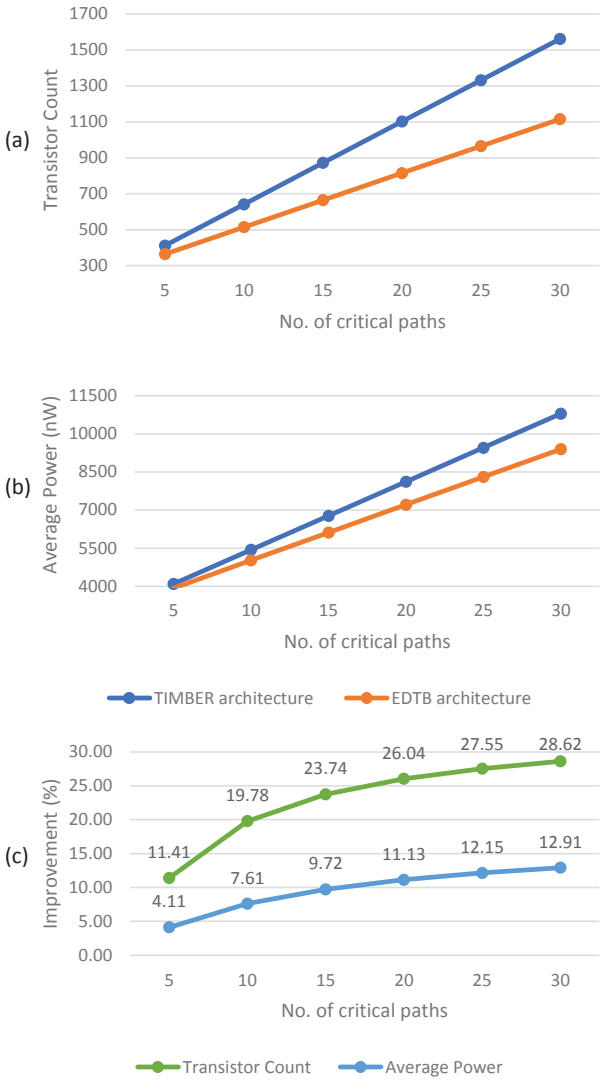


Fig. 6. Comparative analysis for transistor count and avg. power

IX. CONCLUSION

This paper presented a novel area-efficient architecture for online timing error resilience using continuous time-borrowing which is capable of masking multi-stage timing error. Since an additional recovery or replay operation for error management is not required, the proposed architecture can minimize the performance penalty without slowing down the clock speed. As a result, a DVFS system employing the proposed architecture can benefit from reduced safety margins, resulting in increased operating frequency and throughput, even at a high activation probability of critical paths. Hence, the energy efficiency of the system is improved without much compromise in performance. The proposed implementation also shows a remarkable improvement of up to 34% in transistor count, 50% in clock routing, and up to 18% in average power over the previous multi-stage error masking architecture.

REFERENCES

[1] H. Kaul, M. Anders, S. Hsu, A. Agarwal, R. Krishnamurthy and S. Borkar, "Near-threshold voltage (NTV) design — Opportunities and challenges," DAC Design Automation Conference 2012, San Francisco, CA, USA, 2012, pp. 1149-1154.

[2] S. Das et al., "RazorII: In Situ Error Detection and Correction for PVT and SER Tolerance," in IEEE Journal of Solid-State Circuits, vol. 44, no. 1, pp. 32-48, Jan. 2009, doi: 10.1109/JSSC.2008.2007145.

[3] K. A. Bowman et al., "Energy-Efficient and Metastability-Immune Resilient Circuits for Dynamic Variation Tolerance," in IEEE Journal of Solid-State Circuits, vol. 44, no. 1, pp. 49-63, Jan. 2009, doi: 10.1109/JSSC.2008.2007148.

[4] M. R. Choudhury, V. Chandra, R. C. Aitken and K. Mohanram, "Time-Borrowing Circuit sDesigns and Hardware Prototyping for Timing Error Resilience," in IEEE Transactions on Computers, vol. 63, no. 2, pp. 497-509, Feb. 2014, doi: 10.1109/TC.2012.190.

[5] Clark, Lawrence & Vashishtha, Vinay & Shifren, Lucian & Gujja, Aditya & Sinha, Saurabh & Cline, Brian & Ramamurthy, Chandarasekaran & Yeric, Greg. (2016). ASAP7: A 7-nm finFET predictive process design kit. Microelectronics Journal. 53. 105-115. 10.1016/j.mejo.2016.04.006.

[6] K. Chae, S. Mukhopadhyay, Chang-Ho Lee and J. Laskar, "A dynamic timing control technique utilizing time borrowing and clock stretching," IEEE Custom Integrated Circuits Conference 2010, San Jose, CA, USA, 2010, pp. 1-4, doi: 10.1109/CICC.2010.5617392.

[7] M. Ahmadi, B. Alizadeh and B. Forouzandeh, "A Hybrid Time Borrowing Technique to Improve the Performance of Digital Circuits in the Presence of Variations," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 64, no. 1, pp. 100-110, Jan. 2017, doi: 10.1109/TCSI.2016.2600750.

[8] S. Agwa, E. Yahya and Y. Ismail, "ERSUT: A Self-Healing Architecture for Mitigating PVT Variations Without Pipeline Flushing," in IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 63, no. 11, pp. 1069-1073, Nov. 2016, doi: 10.1109/TCSII.2016.2548261.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:00:05 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2021.3103851, IEEE Transactions on Power Electronics

Time-Domain Characterization and Detection of Open Circuit Faults for H-Bridge Power Cell

Mayank Kumar, *Senior Member, IEEE*

***Abstract*— This paper presents time-domain characterization of semiconductor open circuit switch faults (OCSF's) such as open switch with healthy diode and switch-diode pair open for H-bridge power cell. The inverter fault signatures under diagonally opposite open switch fault conditions with healthy diode are approximately similar therefore this paper presents a detection algorithm with unique identification of each faulty switch within H-bridge power cell. The proposed algorithm is also designed to detect and identify multiple OCSF's. The non-ideal characteristics of semiconductor switches are considered for the identification of switch faults. The well-known unipolar sinusoidal pulse width modulation technique is used for the switching of semiconductor devices for single power cell. The mathematical expressions have also been derived to find out the effect of switch fault using double Fourier integral solution and the switching states. The analytical expressions have been developed for the switched output voltage waveforms under ideal switching conditions. The behavior of the switch faults has been analyzed using inverter output voltage, load current and total harmonic distortion (THD) of pulse width modulated inverter (PWMI) output voltage.**

***Index Terms*—cascaded H-bridge, double Fourier series, fault detection, fault identification, harmonic distortion, open switch faults, parasitic capacitance.**

I. Introduction

THE probability of fault occurrence in the power electronics system increases with the increased number of power semiconductor devices and therefore reliability of the system decreases. The optimal productivity with improved output qualities and operational safety are the primary conditions for industries and production tools. It covers the fault detection in a short time period to avoid hazards. The fault detection system is not only reduce the maintenance cost but also prevents the partial or fully shut-down of the complete system. Thus, the optimal productivity can be achieved by employing fault detection system. According to the survey and research applied on 80 companies and more than 200 industrial processes, 21% of faults detected in power semiconductor devices and 13% in solders. In a total 34 % of faults in a converter is due to failure of device module. The rest of the faults are divided into dc link capacitor faults, load faults and gate driver faults [1], [2].

Multilevel inverters (MLI's) are the most promising structure in the area of medium and high-power applications and for every industrial process with reduced operational cost and quality production [3]. Several basic structures of MLI's are available in the literature such as: neutral point clamped, flying capacitor and cascaded H-bridge (CHB) [4], [5]. The CHB inverters are one of the most popular and promising MLI topology, where number of series connected power cells are used to produce the stepped inverter voltage output. There are several applications of CHB MLI such as renewable energy systems, AC motor drives, reactive power compensation, hybrid energy systems; where fault tolerant capabilities improve power quality and operational hours.

Manuscript received July 31, 2020; revised March 16, 2021 and June 15, 2021; accepted August 08, 2021.

The author is with the Department of Electrical Engineering, Delhi Technological University, New Delhi 110042, India (e-mail: mayankkumar@dtu.ac.in).

The possibilities of faults occurred in a power semiconductor device are due to external and internal causes. The external possibilities of switch failure may be swell or dip from the external supply connected to the terminals of MLI's, high dynamic changes at the load side, short circuit of the load etc. On the other side, the internal possibilities of switch failures may be thermo-mechanical fatigue, gate misfiring, saturation of semiconductor materials etc. Due to the above-mentioned causes, the power switches may be either permanently opened or closed. This phenomenon is defined as fault in the switching device. If switch is permanently open then it is defined as open circuit switch faults (OCSF's) and if the switch is permanently closed then it is defined as short circuit switch faults (SCSF's) [6]. The anti-parallel diode across the switching device has also possibilities of open circuit or short circuit fault.

In literature, several fault tolerant controlled operations of medium and high voltage converters are reported in [7], [8], especially for CHB MLI in [9], [10], and modular multilevel converters with phase shifted carrier (PSC) modulation technique in [11]. The SCSF's are severe in nature and it can cause immediate damage to the complete system. Therefore, fast detection and isolation / protection circuitry are required. Fast acting fuses or fast acting gate drive circuit protections are generally used for the protection purpose. The SCSF should be clear within t_{sc} ($\approx 10\mu$sec), where t_{sc} depends on semiconductor chip, gate driver voltage i.e., reduction in gate-emitter voltage V_{GE} reduces the fault current magnitude resulting increase in short-circuit time t_{sc} [12]. In OCSF conditions, the MLI can operate under faulty condition with reduced output quality because of inherent switching redundancy. However, this may lead to increased voltage stress across the other healthy switches and it may be the cause of complete system damage. Therefore, to make the system fault tolerant, the fault must be detect and the fault location must be identified in minimum time duration.

In literature, several OCSF detection and identification techniques are reported. These techniques are classified in general as (i) hardware-based fault detection methods (H-FDM) and (ii) computational-based fault detection methods (C-FDM) [13], [14]. The hardware-based techniques are used to detect severe faults such as short-circuit faults in modules and shoot-through of legs. Such faults can be isolated using advanced inbuilt gate driver protection circuits, where fault isolation is required within 10μsec [12], [15]. The fault detection methods based on gate signal monitoring is used to detect SCSF's, OCSF's; it is presented with additional voltage / current sensors and the measured quantities are compared with defined threshold values in [16], [17].

On the other hand, for non-sever faults such as C-FDM's are used, it requires additional processing of measurement data. Such computational methods use either an extracted feature like switching frequency component of the module or the output states of the system to detect and localize the faults. The C-FDM's can be further categorized as model-based fault detection

0885-8993 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:44:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2021.3103851, IEEE Transactions on Power Electronics

methods (M-FDM's) and feature-based fault detection methods (F-FDM's). The main advantage of model-based fault detection technique is to minimize the inclusion of additional sensors needed for the measurement of featured quantities. The occurrence of OCSF in any active switch of an H-bridge affects the inverter output voltage. Several M-FDM's are available in the literature, where high bandwidth voltage sensors are used to measure either the inverter output voltage across each H-bridge module [18] or CHB MLI per phase output voltage [13], [19], [20]. Since for the detection of OCSF's, switching states are used for the comparison of respective inverter output voltage and load current therefore M-FDM's are much faster. The mathematical modelling of switch faults diagnosis in multilevel-inverter drive (MLID) systems are very difficult due to many switching devices and their nonlinear effects, therefore an artificial intelligence and fast Fourier transform with neural network-based fault diagnostic method for CHB MLID is presented in [21]. Several literatures are based on model predictive control algorithm, which predicts the future states and used to compare against the actual state for fault detection [22], [23]. The accuracy of diagnosis depends on the predicted values of future states and if the parameters are inaccurate with respect to real time, the probability of diagnostic accuracy reduces. The extracted characteristics of the system outputs are used to detect the fault under F-FDM's. Switching frequency based harmonic component analysis is one of the popular characteristics of the F-FDM's [24], [25]. In harmonic component analysis, the threshold level is categorized in dc components, side-band components and base-band components. Similar to harmonic component based open circuit fault detection method few other F-FDM's are available in the literature such as: electromagnetic interference (EMI) based and conduction based [26], [27].

In this paper, fast detection algorithms for single H-bridge power cell are proposed and detection logic designs have been developed to minimize the detection time and identify the accurate switch fault locations for different type of OCSF's. Three different open switch fault categories are considered for the analysis: *(i)* switch open circuit faults with healthy diode $SF_{xy\text{soc}}$ (where subscript *x*, *y* represents the upper/lower switch and *leg* number in the power cell and 'soc' represents switch open circuit with healthy diode); *(ii)* diode open circuit faults with healthy switch $SF_{xy\text{DOC}}$ (where 'DOC' represents diode open circuit with healthy switch); *(iii)* open circuit faults $SF_{xy\text{oc}}$ (where 'oc' represents switch-diode pair is under fault). In H-bridge power cell, the fault signatures of diagonally opposite switch faults are approximately similar therefore it is difficult to identify unique switch fault location. Therefore, a unique fault detection algorithm is developed for identification of individual switches within the power cell. The proposed switch fault detection and identification technique is also able to detect multiple open switch faults. The role of non-ideal behavior of semiconductor switches are used for the identification of switch faults. The phase-shifted pulse width modulation (PSPWM) technique is used to generate switching logics for semiconductor devices. The primary contributions of this paper are: 1) The detection logic designs are developed for OCSF's; 2) fault appeared in any of the switch within H-bridge power cell can be uniquely identified within couple of switching intervals; 3) the proposed detection logic designs are presented for rapid detection and identification; 4) the designed algorithm can detect multiple open switch faults; 5) mathematical relations are developed under OCSF's conditions; 6) experimental results are also presented for the verification of proposed algorithm.

II. IGBT Switching Characteristics

The switching behavior of an IGBT depends on the structure, inner / outer resistances, and internal capacitances. The half bridge structure with an input gate resistance R_{GI} and internal capacitances C_{GC}, C_{GE} is presented in Fig. 1 (a). The voltage dependent low signal capacitances play major role during turn-on and turn-off of the transistors. These voltage dependent parasitic capacitances are classified as input capacitance (C_{ies} = C_{GC} + C_{GE}), reverse transfer capacitance (C_{res} = C_{GC}), and output capacitance (C_{oes} = C_{GC} + C_{CE}); where C_{GC}, C_{GE}, C_{CE} are internal capacitances between gate-collector, gate-emitter, and collector-emitter, respectively [28]. The collector-emitter capacitance is considered equivalent to the anti-parallel diode capacitance.

The equivalent simplified model of the IGBT switch circuit seen from the gate to the emitter is presented in Fig. 1 (b). The circuit model is developed without consideration of base resistor small modulation voltage and the small turn-on voltage of the diode in the internal bipolar junction transistor for IGBT. Therefore collector–emitter voltage V_{CE} is equivalent to drain–source voltage V_{DS}. The gate to drain region is of major concern because of higher mobility of carriers in the channel during turn-on process, the capacitance of this region C_{GD} can be modelled using variable depletion zone capacitance C_{GDJ} and fixed oxide zone capacitance C_{OXD} [29]. The turn-on and turn-off characteristics of transistor are shown in Fig. 1 (c) and (d), respectively. The explanation of all the regions is readily available in the literature [2], [30]. The switching logics of both the switches in the same *leg* are complementary in nature. Further, if the upper switch Sw_{u1} is switched-off then the voltage developed across the switch is V_{dc}. To make the switch Sw_{u1} turned-on, the gate charge is needed and it is fulfilled by gate current I_G. Similarly, total gate charge required for turn-on and turn-off process of the transistor is supplied from driver circuit.

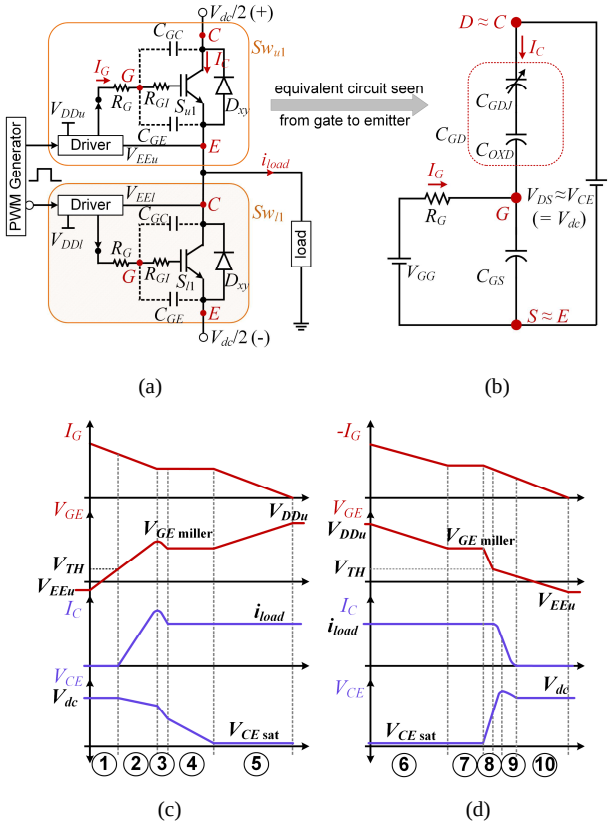


Fig. 1 Switching characteristics of IGBT (a) circuit diagram of single PWM inverter *leg* (b) equivalent switch circuit model (c) turn-on and (d) turn-off behavior.

0885-8993 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:44:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2021.3103851, IEEE Transactions on Power Electronics

The effect of non-idealities of semiconductor switches are used during turn-on and turn-off to identify the specific switch fault location. According to the behavior of output voltage, current characteristics and switching logics, the unique identification technique is developed for OCSF's. Figs. 2 (a) and (b) show the current path during fault in S_{u1} with healthy diode (SF_{u1SOC}) and fault in S_{l2} with healthy diode (SF_{l2SOC}) for a three-level H-bridge or a power cell of CHB inverter; respectively. The two legs (i.e., *leg* 1 and *leg* 2) of a power cell consist: two upper switches S_{u1}, S_{u2}; and two lower switches S_{l1}, S_{l2} with anti-parallel diodes D_{u1}, D_{u2} and D_{l1}, D_{l2}, respectively. The combination of switch-diode pair is represented as Sw_{xy}, where x is upper/lower switch and y is *leg* number. The switching behavior of an international rectifier AUIRGP4063D (IGBT with ultrafast soft recovery diode) is captured using MultiSim software and it is plotted in Figs. 2 (c) and (d) under SF_{u1SOC} and SF_{l2SOC}; respectively. Figs. 2 (c) and (d) show the voltages across the switches in the respective faulty *leg*, load current $i_l|_{soc}$ and pulsewidth modulated inverter (PWMI) output voltage $v_o|_{soc}$. These cases are possible only with healthy diode operation where freewheeling action of current is possible. The switching behavior under SF_{u1SOC} can be discussed for three different intervals as follows:

T_1: S_{u1} - under fault and S_{u2} - ON: the switch voltages $V_{Su1} \approx 0$ (small on-state diode voltage), $V_{Sl1} = V_{dc}$ and $v_o|_{soc} \approx 0$; current direction is negative and load current path is shown in Fig. 2 (a), i.e. from D_{u1} to S_{u2} to load;

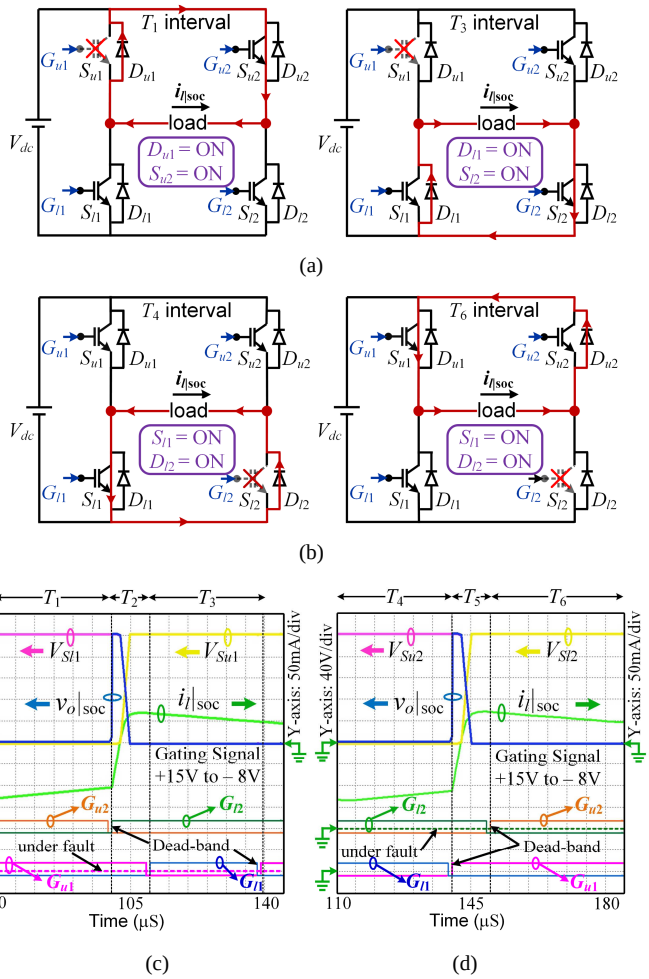


Fig. 2 Characterization of switch faults (a), (b) H-bridge cell with load current path under SF_{u1SOC}, SF_{l2SOC}, respectively; and (c), (d) switch voltages across faulty *leg*, output voltage, load current, and switching states under SF_{u1SOC}, SF_{l2SOC}, respectively.

T_2: transition state (TS): direction of current changes from negative to positive therefore anti-parallel diode D_{l1} is switched-on and the voltage developed across V_{Sl1} is released through load; load current path is changed from T_1 interval to T_3 interval;

T_3: S_{l1} - ON and S_{l2} – ON: the switch voltages $V_{Sl1} \approx 0$, $V_{Su1} = V_{dc}$ and $v_o|_{soc} \approx 0$; current direction is positive and load current path is shown in Fig. 2 (a), i.e., from D_{l1} to load to S_{l2}.

Similarly, the operation can be discussed for diagonally opposite switch fault SF_{l2SOC} as follows:

T_4: S_{l2} - under fault and S_{l1} - ON: the switch voltages $V_{Sl2} \approx 0$ (small on-state diode voltage), $V_{Su2} = V_{dc}$ and $v_o|_{soc} \approx 0$; current direction is negative and load current path is is shown in Fig. 2 (b), i.e., S_{l1} to D_{l2} to load;

T_5: transition state (TS): direction of current changes from negative to positive therefore anti-parallel diode D_{u1} is switched-on and the voltage developed across V_{Su1} is released through load; load current path is changed from T_4 interval to T_6 interval;

T_6: S_{u1} - ON and S_{u2} – ON: the switch voltages $V_{Su2} \approx 0$, $V_{Sl2} = V_{dc}$ and $v_o|_{soc} \approx 0$; current direction is positive and load current path is shown in Fig. 2 (b), i.e., from S_{u1} to load to D_{u2}.

III. Characterization and Identification of Open Switch Faults in H-Bridge Power Cell's

The healthy operation of the H-bridge power cell output with inductive load can be classified into four operating modes with respect to PWMI voltage output and load current characteristics as given in Table I, where v_{mod1} is sinusoidal modulating waveform with zero phase shift. All the four operating modes are possible only under healthy operation with inductive load. Therefore, the detection technique is developed with respect to PWMI output voltage error v_{err} with respective switching states, load current i_l; where PWMI output voltage error v_{err} is the unit resultant of actual PWMI output voltage to the reference PWMI output voltage.

Fig. 3 (a) shows well known unipolar sinusoidal pulsewidth modulation (SPWM) technique, where the gating signals are the resultant of comparison between two 180° phase-shifted modulating sinusoidal signals $v_{mod1}(t)$, $v_{mod2}(t)$ and a high frequency triangular carrier signal $v_{tri}(t)$. The three-level voltage output (i.e., $V_o = +V_{dc}$, 0, $-V_{dc}$) with switching states and load current under healthy operation of H-bridge power cell is shown in Fig. 3 (b). The generated gating logics G_{u1}, G_{l2} and G_{u2}, G_{l1} are shown in Figs. 3 (c) and (f), respectively. The PWMI voltage output with all switching states and load current of the H-bridge power cell under SF_{u1soc}, SF_{l2soc}, SF_{u2soc}, and SF_{l1soc} are shown in Fig. 3 (d), (e), (g) and (h); respectively. The instantaneous switching states over a fundamental period under PSPWM technique is presented. Due to the space constraint the switch and diode symbols are changed in Fig. 3 as follows: the switches S_{u1}, S_{l1}, S_{u2}, S_{l2} are replaced by S_1, S_2, S_3, S_4, respectively and the diodes D_{u1}, D_{l1}, D_{u2}, D_{l2} are replaced by D_1, D_2, D_3, D_4, respectively. The similar switching behavior is achieved under same modes of operation in all fault conditions but a unique fault mode is appeared with different SOC faults. In TS, the load current direction i_l changed with respect to the direction of v_{mod1}. The PWMI voltage output or the fault signature under diagonally opposite switch pairs SF_{u1soc}, SF_{l2soc} and SF_{u2soc}, SF_{l1soc} are similar therefore a unique detection algorithm is designed for identification of faulty switch within H-bridge power cell. The identification and characterization of three different types of faults are discussed in further sub-sections.

0885-8993 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:44:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2021.3103851, IEEE Transactions on Power Electronics

TABLE I
OPERATING MODES

Operating modes	mode 1	mode 2	mode 3	mode 4
v_{mod1}	positive	positive	negative	negative
i_l	positive	negative	positive	negative

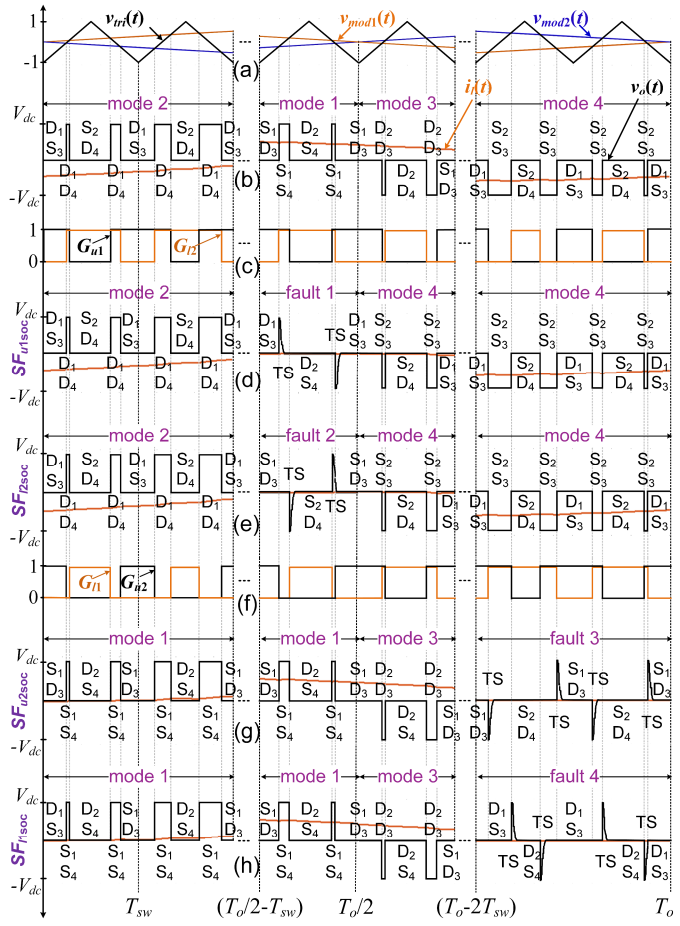


Fig. 3 Switching characteristics and modes of operations of power cell under different switch fault conditions; (a) comparison between triangular carrier and modulating signals, (b) PWMI output voltage and load current under healthy condition, (c) switching signal G_{u1}, G_{l2}, (d), (e) PWMI output voltage and load current under SF_{u1soc}, SF_{l2soc}, respectively, (f) switching signal G_{u2}, G_{l1}, (g), (h) PWMI output voltage and load current under SF_{u2soc}, SF_{l1soc}, respectively.

A. *Open Switch Fault Identification with Healthy Diode* (SF_{xysoc})

Figs. 4 (a) and (b) show the output of H-bridge power cell under SF_{u1soc} and SF_{l2soc}, respectively. It can be seen from the zoomed section that the direction of current changes during transition states under fault condition. During SF_{u1soc} fault when the value of $i_l|_{soc} > 0$ for switching logic G_{l2} is high and $i_l|_{soc} < 0$ for switching logic G_{l2} is low. Similarly, under other switch fault conditions, the switching states and load current relation can be identified. The relation between voltage spikes and switching logics can also be developed with respect to Figs. 4 (a) and (b) for SF_{u1soc} and SF_{l2soc}; respectively. It can be seen from the zoomed section of Fig. 4 (a) that a negative and a positive voltage spikes is appeared with the rising edge of G_{u2} and G_{l2}, respectively, when $v_{mod1} > 0$ and $i_l|_{soc} \approx 0$. The zoomed section of Fig. 4 (b) shows that a negative and a positive voltage spikes is appeared with the rising edge of G_{l1} and G_{u1}, respectively, when $v_{mod1} > 0$ and $i_l|_{soc} \approx 0$. Similarly, the relation can be developed between voltage spikes and switching logics for other diagonally opposite switch faults SF_{u2soc} and SF_{l1soc}.

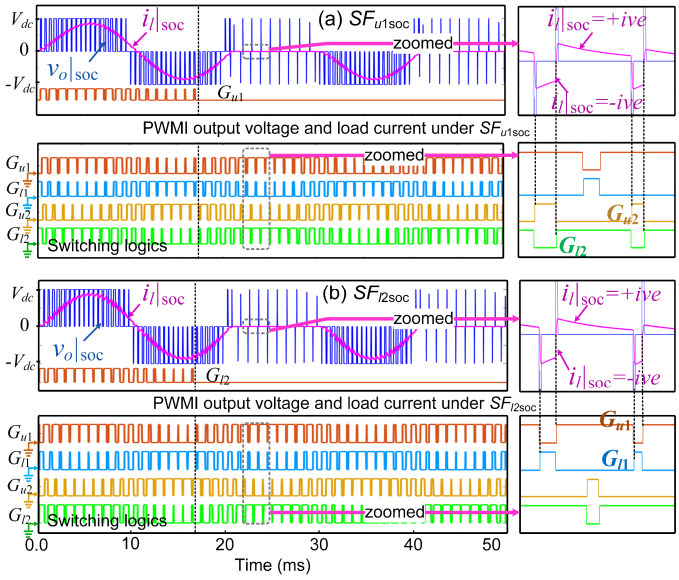


Fig. 4 The load current, PWM output voltage, switching states and the zoomed view under (a) SF_{u1soc}, and (b) SF_{l2soc}.

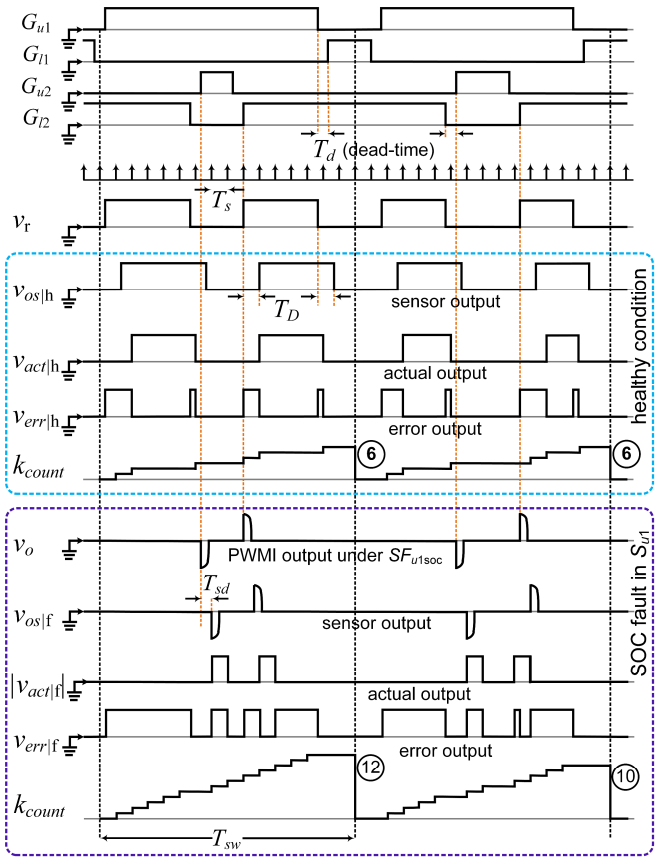


Fig. 5 Characterization of PWMI output voltage under healthy, SF_{u1soc} conditions in mode 1 (i.e., positive voltage and current) and development of fault detection techniques with the consideration of different system delays.

Fig. 5 shows the unit magnitude of PWMI output voltage and designing procedure of fault detection techniques with the consideration of different system delays under healthy and SF_{u1soc} conditions during mode 1 i.e., $v_{mod1} > 0$ and $i_l > 0$. The unit magnitude of ideal inverter output voltage can be defined as the summing resultant of G_{u1} and $-G_{u2}$; further, the rising and falling edges of the resultant are shifted with dead-time T_d and it is defined as reference PWMI output voltage v_r. The voltage sensor outputs are presented as $v_{os|h}$ and $v_{os|f}$ under healthy and faulty conditions, respectively. The combined effect of sensor

0885-8993 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:44:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2021.3103851, IEEE Transactions on Power Electronics

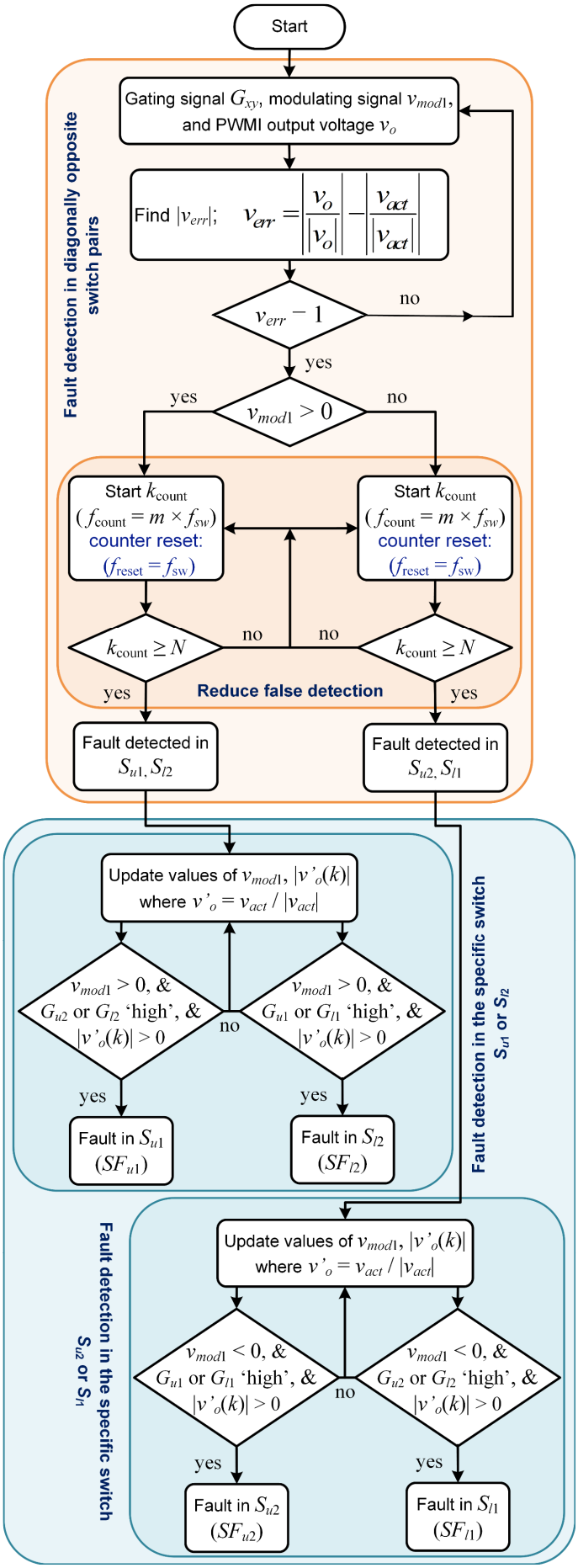


Fig. 6. Flowchart for the proposed OCSF algorithm with fault detection and fault identification.

delay (T_{sd}), processing delay of the controller and turn-on / turn-off switching delays are considered as T_D. Further with respect to the sampling time (T_s) of the analog to digital (A/D) converter, the voltage which is compared with reference PWMI output voltage v_r is defined as actual PWMI output voltage $v_{act|h}$ and $v_{act|f}$ under healthy and faulty conditions, respectively. The counter frequency to switching frequency ratio m is considered as 16. Fig. 5 shows that the consideration of non-ideal parameter of system delays and it is responsible for the non-zero values of $v_{err|h}$ and resulting the counter value reaches at k_{count} = 6. Since the value of k_{count} = 6 (= $k_{healthy}$, i.e., the value of counter under healthy condition), therefore, to reduce the probability of false detection, the value of k_{count} should be greater than $k_{healthy}$. Whereas during SOC fault, the non-zero instants of $v_{err|f}$ increases, resulting the counter value reaches at k_{count} = 12. Since the value of $k_{count} > k_{healthy}$ i.e., greater than the counter value under healthy condition, therefore the fault detection logic signal becomes 'high'.

Fig. 6 shows the flowchart of the proposed OCSF detection and identification algorithm. The flowchart is divided into two parts and each part is further divided into two sub-parts. The first part (under orange shade) represents the procedure to detect the fault under diagonally opposite switch pairs and sub-part is used to reduce the probability of false detection. The counter frequency f_{count} is equal to ($m \times f_{sw}$); where m is an integer and f_{sw} is the switching frequency. Threshold value N is chosen, such that $k_{healthy} < N < m$. With the increase in the value of N from $k_{healthy}$, the probability of false fault identification reduces. It also slighty increases the switch fault identification period. In this manuscript, $N > 0.5 \times m$ is chosen to minimize false detection and in real time implementation m = 10 and N = 7 is considered. The second part (under blue shade) represents the identification of specific switch fault.

The proposed OCSF algorithm is implemented and it is presented as a detection logic design block for SOC faults ($DLDB_{soc}$) in Fig. 7. The specific fault locations SF_{xysoc} can be detected and identify using three input variables such as actual output voltage v_{act} (sensor output), respective switching logic and v_{mod1} (internal signal generated by controller). The relation between switching logics and specific switch faults, the faulty modes and partial faulty modes under different switch faults are mentioned in Table II. The internal schematic of $DLDB_{soc}$ is presented in Fig. 7 (b). It is divided into three parts: *(i)* the fault signal is bifurcated into diagonally opposite switch pairs; *(ii)* to reduce the probability of false fault detection; *(iii)* the fault signal is bifurcated within diagonally opposite switch pairs, i.e., the specific switch fault locations are identified. The second and third parts of design logic is similar for all the four switch faults. To generate constant frequency pulse train, a high frequency (= $m \times f_{sw}$) triangular waveform is used. In the simulation results the value of m = 100 is considered, whereas the value of N = 70 is chosen. If a continuous high logic output is generated through an AND logic operator at C1 within switching period T_{sw} then the counter generates a ramp with peak magnitude m and frequency f_{sw}. It is further compared with a constant value N. If the magnitude of generated ramp signal is higher than the constant value then the fault detection logic becomes high and one of the switch pair SF_{u1soc}, SF_{l2soc} or SF_{u2soc}, SF_{l1soc} is detected within a couple of switching frequencies. Further the output is processed through Sys-B, through which the exact switch fault location for SOC fault is identified. Once the fault is detected through the $DLDB_{soc}$ then the system should be reset for the next fault detection and switch identification.

0885-8993 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:44:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2021.3103851, IEEE Transactions on Power Electronics

TABLE II
FAULT DETECTION SIGNAL UNDER DIFFERENT OCSF'S CONDITIONS

Fault location	SF_{u1soc}	SF_{l1soc}	SF_{u2soc}	SF_{l2soc}
Detection signal	G_{l2}	G_{u2}	G_{l1}	G_{u1}
Faulty mode	mode 1	mode 4	mode 4	mode 1,
Partial faulty mode	mode 3	mode 2	mode 2	mode 3

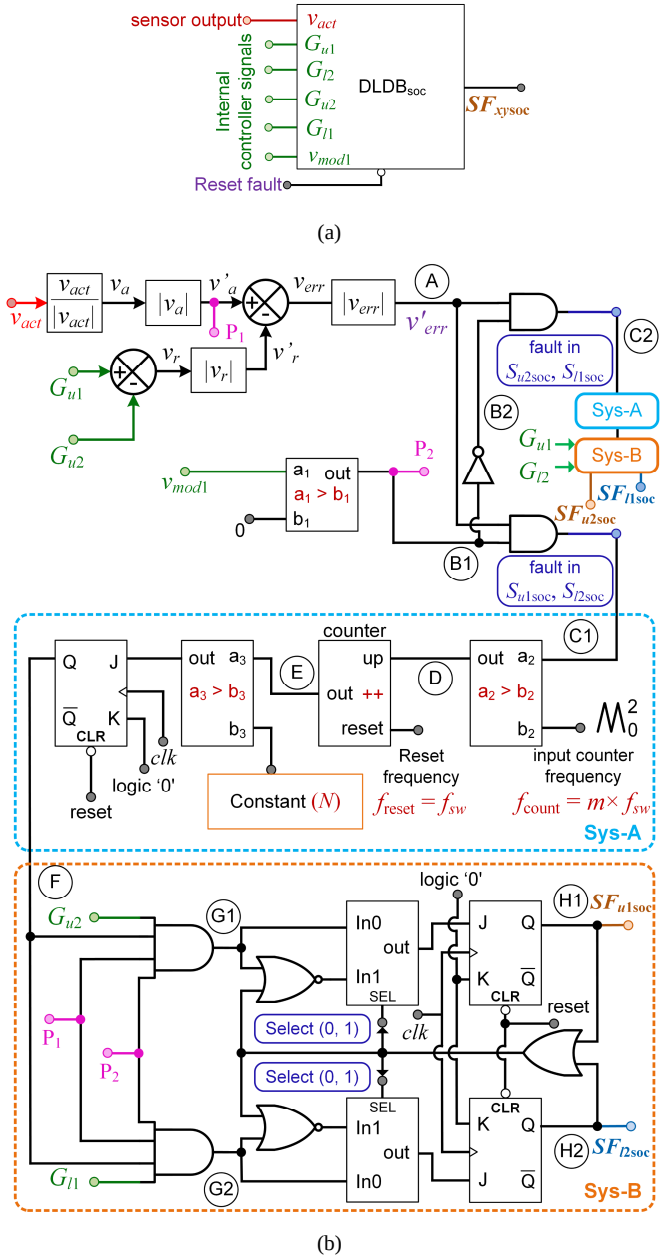


Fig. 7 Proposed SOC fault detection logic design; (a) block diagram with respective input and output fault logic, (b) internal schematic digram for fault identification system.

B. Open Diode Fault Identification with Healthy Switch (SF_{xyDOC})

The probability of SF_{xyDOC} is very low and uncommon. This type of fault can be detected using three input variables such as load current i_l, modulating signal v_{mod1} and respective switching logic. The detection logic design block for DOC faults (DLDB$_{DOC}$) and internal schematic of DLDB$_{DOC}$ are shown in Fig. 8 (a) and (b), respectively. The relationship between input switching logics and output diode fault location is also mentioned in Fig. 8. The counter frequency is m times of the switching frequency. Since the load current becomes zero within two switching period due to unavailability of freewheeling

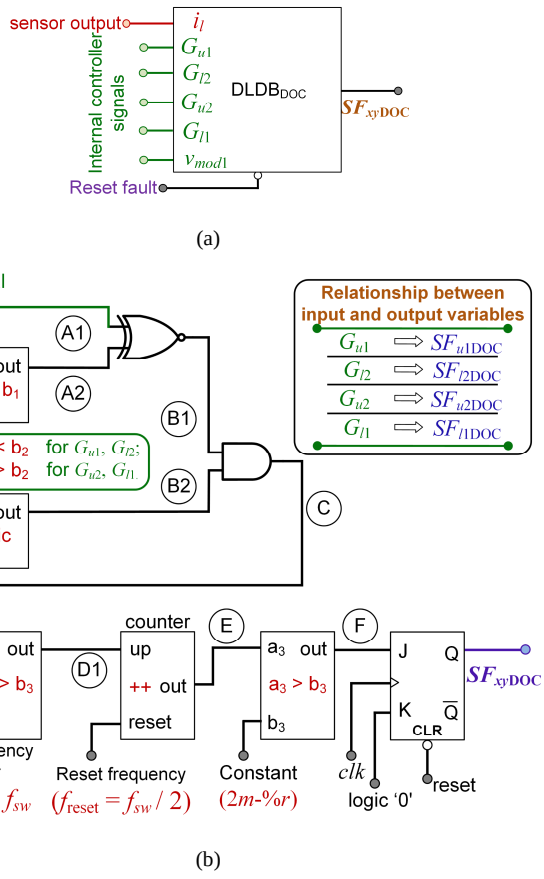


Fig. 8 Proposed DOC fault detection logic design; (a) block diagram with respective input and output fault logic, (b) internal schematic diagram of fault identification system.

current path under open diode fault therefore the counter is reseted after two switching periods. The counter is compared with a constant value (= 2m - %r), where r represents the system delays. The value of r is chosen 4% of the value of m.

C. Switch-Diode Pair Open Fault Identification (SF_{xyoc})

The free-wheeling action of current through diode under SF_{xyoc} is not possible resulting huge voltage spike during switching in the power cell similar to the DOC fault. The healthy PWMI output voltage is also missing similar to the SOC fault. This type of fault can be detected using AND operation of the output of DLDB$_{SOC}$ and DLDB$_{DOC}$. If the output of these two-detection block signal logic '*high*' indicates the fault in the respective switch-diode pair.

IV. FORMULATION AND ANALYSIS OF OCSF'S

The analytical solutions of OCSF's in H-bridge power cell with PSPWM technique can be mathematically derived in time-domain using double Fourier integral with the consideration of ideal switching conditions. The integral solutions are obtained using Jacobi-Anger expansions [31], [32]. The three-level inverter output voltage can be derived using (1), where $u_a(t)$, and $u_b(t)$ are the analytical expressions of output gating signal generated using PSPWM with phase shifted modulating signals at $\theta = 0$ and $-\pi$ for *leg* 1 and *leg* 2, respectively. The expression for u_a and u_b are available in the literature [33], [34] and it is presented in (2a) and (2b), respectively; where M is modulation index; ω_o, ω_c are fundamental and carrier frequency in rad/sec; θ_c is the phase shift for the carrier signal; $J_p(\xi)$ is the pth order Bessel function; and p, q are base band and side band indices, respectively.

0885-8993 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:44:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2021.3103851, IEEE Transactions on Power Electronics

$$v_o(t)=\left\{u_a(t)\big|_{\theta=0}-u_b(t)\big|_{\theta=-\pi}\right\}V_{dc}=\{u_a-u_b\}V_{dc} \tag{1}$$

$$u_a=\frac{(1+M\cos\omega_o t)}{2}+\frac{2}{\pi}\sum_{p=1}^{\infty}\sum_{q=-\infty}^{\infty}\frac{1}{p}\left\{\sin\left(\frac{(p+q)}{2/\pi}\right)\times J_q\left(\frac{p\pi M}{2}\right)\times\cos\left(p(\omega_c t+\theta_c)+q\omega_o t\right)\right\} \tag{2a}$$

$$u_b=\frac{1+M\cos(\omega_o t-\pi)}{2}+\frac{2}{\pi}\sum_{p=1}^{\infty}\sum_{q=-\infty}^{\infty}\frac{1}{p}\left\{\sin\left(\frac{(p+q)}{2/\pi}\right)\times J_q\left(\frac{p\pi M}{2}\right)\times\cos\left(p(\omega_c t+\theta_c)+q(\omega_o t-\pi)\right)\right\} \tag{2b}$$

The PSPWM inverter output voltage $v_o(t)$ (where $v_o(t) = u_o(t) \times V_{dc}$) and load current i_l under healthy operation can be derived using (2a) and (2b) as follows:

$$u_o(t)=M\cos(\omega_o t)+\frac{4}{\pi}\sum_{p=1}^{\infty}\sum_{q=-\infty}^{\infty}\frac{1}{2p}\times J_{2q-1}(p\pi M)\times\cos\left((p+q-1)\pi\right)\times\cos\left(2p(\omega_c t+\theta_c)+(2q-1)\omega_o t\right) \tag{3}$$

$$u_o V_{dc}-L\frac{di_l}{dt}-Ri_l=0\Rightarrow I_l=\frac{V_{dc}}{sL+R}U_o(s) \tag{4}$$

The PWMI output voltage during SOC and OC are similar and it is derived further; whereas under DOC, the PWMI output voltage is similar as healthy operation as given in (3). As discussed in the previous section, some of the modes of the operation are disappeared or affected during SF_{xysoc}. Mathematical relations are developed for PWMI output voltage $v_o|_{xysoc}$ and load current $i_l|_{xysoc}$ and it is presented in (5); whereas different switch fault relations are derived and presented in (6a) - (6d). The first term of (6) represents the unit magnitude of the PWMI output voltage with the consideration of respective faulty modes as given in Table III. Whereas the second term of (6) represents the characteristics of fault occurred in SF_{u1soc}, SF_{l2soc} and SF_{u2soc}, SF_{l1soc} during $i_l|_{xysoc} > 0$ and $i_l|_{xysoc} < 0$, respectively. It considers the respective partially faulty modes as given in Table III, i.e., the time of fault occurrence follow the current conditions, resulting the change occurred in H-bridge inverter switched output voltage.

$$v_o|_{xysoc}=u_o|_{xysoc}\times V_{dc} \tag{5a}$$

$$sLI_l|_{xysoc}-Li_l|_{xysoc}(0)+RI_l|_{xysoc}=V_{dc}U_o(s) \tag{5b}$$

$$u_o|_{u1soc}=\underbrace{u_o(t)}_{for\ all\ i_l|_{xysoc}\le 0}-\underbrace{u_b(t)}_{for\ all\ i_l|_{xysoc}>0} \tag{6a}$$

$$u_o|_{l1soc}=\underbrace{u_o(t)}_{for\ all\ i_l|_{xysoc}\ge 0}+\underbrace{\overline{u_b(t)}}_{for\ all\ i_l|_{xysoc}<0} \tag{6b}$$

$$u_o|_{u2soc}=\underbrace{u_o(t)}_{for\ all\ i_l|_{xysoc}\ge 0}+\underbrace{u_a(t)}_{for\ all\ i_l|_{xysoc}<0} \tag{6c}$$

$$u_o|_{l2soc}=\underbrace{u_o(t)}_{for\ all\ i_l|_{xysoc}\le 0}-\underbrace{\overline{u_a(t)}}_{for\ all\ i_l|_{xysoc}>0} \tag{6d}$$

V. Results and Discussion

The system parameters are taken as follows: dc voltage supply for single power cell V_{dc} = 40V; fundamental (line) frequency f_o = 50 Hz; carrier (switching) frequency f_{sw} = 1 kHz; modulation index M = 0.9; and load impedance Z_L = 1 Ω, 4 mH, only for the representation of each mode, higher value of L/R is considered and results are presented in Figs. 9 and 10. Further, the algorithm is also verified with other load impedance (i.e., Z_L = 15 Ω, 10 mH; presented in Figs. 11-14, 16) and switching frequency (i.e., f_{sw} = 2 kHz; presented in Fig. 14). These are only test parameters and the results are equally valid on other system parameters and switching frequencies. Figs. 9 (a) – (d) are plotted using respective mathematical relations derived for healthy and SOC fault conditions. The modulating signal is used as a cosine function of time in Fig. 9 for the verification of proposed detection technique. Since the load current characteristics is the function of PWMI output voltage therefore in analytical output waveform it is difficult to consider the effect of $i_l|_{SFu2soc} < 0$ (i.e., second term of (6.c) or partially faulty mode (*mode* 2) as given in Table II) under SF_{u2soc} condition. Since the PWMI output voltage and current waveform changes under fault therefore the total harmonic distortion of the output voltage THD_V increases. It can be seen from Figs. 9 (b) and (d).

The consideration of partially faulty modes and their effects are presented with the help of Matlab simulations in Fig. 10. The SOC faults are triggered using gating signals in the upper switches (i.e., S_{u1} and S_{u2}) of the H-bridge cell. The fault is triggered at t_1 = 32.4 msec in both the cases under SF_{u1SOC} and SF_{u2SOC}. After $t = t_1$, inverter in both the conditions are operating under *mode* 4 followed by *mode* 2, which is not a faulty mode for SF_{u1SOC}, therefore in the first condition healthy output is generated and fault is appeared after t_4= 38 msec; and at t_5 = 40.3 msec, the fault with specific switch location is identified in couple of switching period. It is presented in Fig. 10 (a) - (c). Whereas for SF_{u2SOC}, *mode* 4 is a faulty mode and *mode* 2 is a partially faulty mode, therefore the duration $t = t_1$ - t_3 switch fault occurred because load current has negative value as discussed

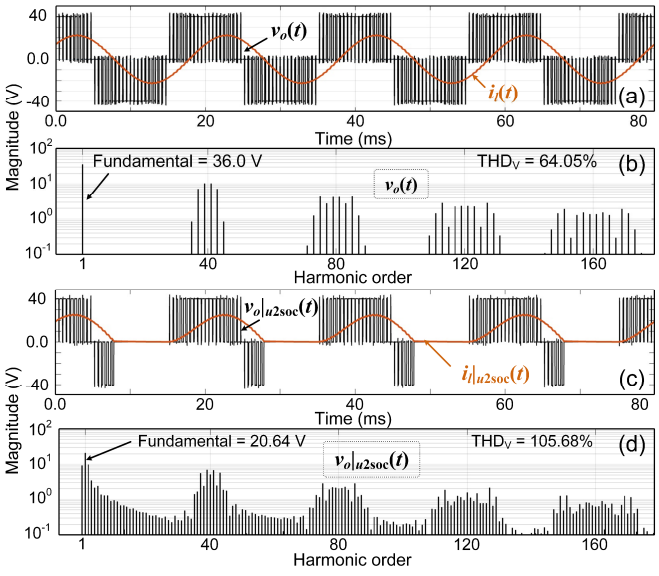


Fig. 9 Analytical results of PWMI output voltage with load current and voltage harmonic spectrum under (a), (b) healthy condition and (c), (d) SF_{u2soc}.

0885-8993 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:44:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2021.3103851, IEEE Transactions on Power Electronics

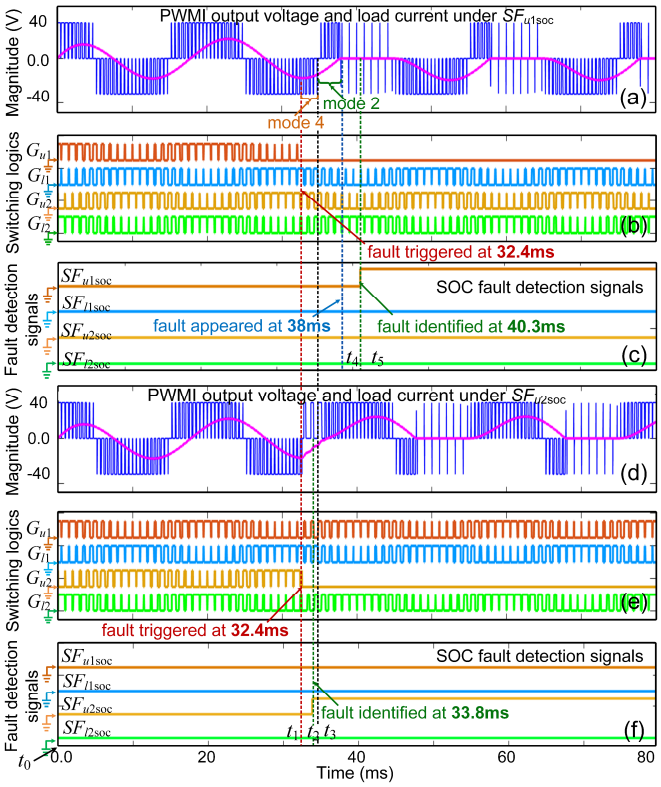


Fig. 10 Characterization, switching states and identification of SOC fault for; (a), (b), (c) S_{u1soc}; and (d), (e), (f) S_{u2soc}; respectively.

in the second term of (6c) and the PWMI output during this interval is $u_a(t)$; thereafter (i.e., $t > t_3$) load current has positive value and the PWMI output during this interval is $u_o(t)$ from the first term of (6c). At $t_2 = 33.8$ msec, the fault with specific switch location is identified after couple of switching period. It is presented in Fig. 10 (d) - (f).

Further, load impedance Z_L = 15 Ω, 10 mH and sinusoidal modulating signal are used for verification of proposed algorithm. The algorithm is verified for multiple switch faults in Fig. 11, which may occur due to over current transients in a H-bridge. It may cause for simultaneous open circuit faults in both the switches of the same leg. Figs 11 (a-c) show the characteristics under single open switch fault in S_{u1oc} at $t_1 = 42$ msec and OCSF is identified at $t_2 = 44.1$ msec. The switch identification time period T_{sid} (i.e., the time duration between fault appeared in the inverter output and faulty switch identification) is similar to Figs 10 (a-c) with different load. Figs 11 (d-f) show the characteristics under multiple open switch faults in the same leg, i.e., in S_{u1oc}, S_{l1oc} at $t_1 = 42$ msec. Similar to the previous case the fault in S_{u1oc} is identified at $t_2 = 44.1$ msec, whereas SF_{l1oc} is appeared at $t_3 = 50.05$ msec and faulty switch is identified at $t_4 = 52.4$ msec. This shows the verification of proposed algorithm with different load and multiple open switch fault conditions.

Fig. 12 shows the detection and identification of switch-diode pair fault in S_{l1oc}. The OC faults occurred because of complete breakdown of switch with anti-parallel diode combination, loosely connected wires within the H-bridge *leg*, etc. For such type of fault generation, circuit breakers are used before each switch-diode pair in the H-bridge. In OC faults, the feedback path of load current is unavailable during respective switching time region therefore the load current becomes zero once in each switching period with single open switch faults. Resulting a large voltage spikes appeared in each switching interval to make the load current zero within the small-time duration. To verify the OC faults experimentally circuit breakers / jumpers can be used with switch-diode pair. Fig. 12 shows all the behavior of switch faults discussed before. The circuit breaker across Sw_{l1} is opened at $t_1 = 21$ msec (i.e., $t_0 - t_1$: healthy operation); the effect of diode fault occurred after $t_2 = 21.3$ msec, thereafter huge voltage spikes appeared at PWMI output and DOC fault detection logic became high within couple of switching period at $t_3 = 23.9$ msec; again, from $t_4 = 31.1$ msec, the effect of switch fault occurred and SOC fault detection logic became high within couple of switching period at $t_5 = 33.5$ msec. The OCSF's (i.e., SF_{l1soc} and SF_{l1oc}) can be identified in couple of switching period and thereafter the post fault control technique can be implemented. From zoomed sections of the respective Fig. 12, it can be seen that the load current became zero when G_{l1} is high under SF_{l1oc}. The similar logics are used to detect OC faults. Since the fault is triggered at $t_1 = 21$msec (i.e., $v_{mod1} > 0$), therefore initially DOC fault is detected and thereafter SOC fault is detected.

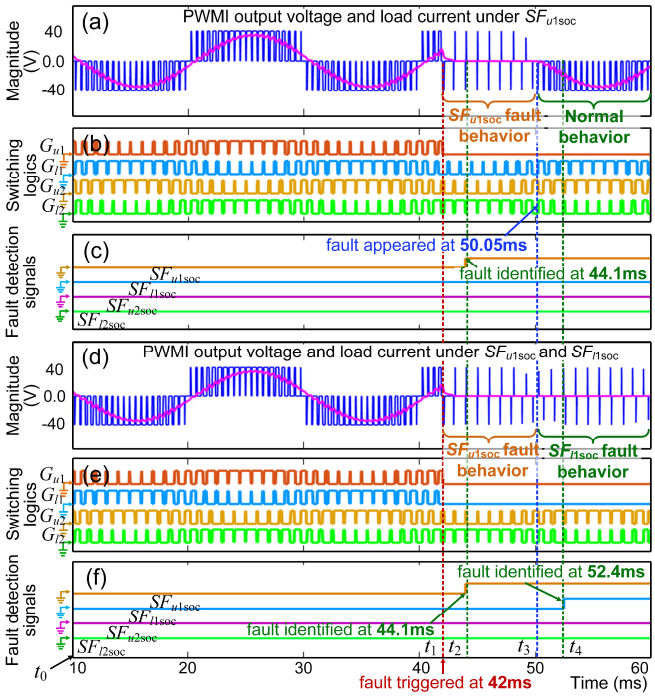


Fig. 11 Characterization, switching states and identification of SOC faults for (a), (b), (c) single switch fault SF_{u1soc} and (d), (e), (f) multiple switch faults in the same leg SF_{u1soc}, SF_{l1soc}; respectively.

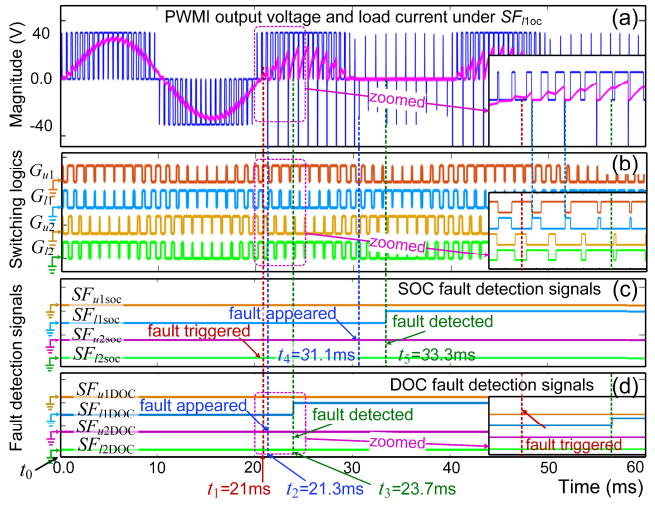


Fig. 12 Detection and identification of OC faults: PWMI output voltage with load current, switching states, and fault detection signals under SF_{l1oc}.

0885-8993 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:44:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2021.3103851, IEEE Transactions on Power Electronics

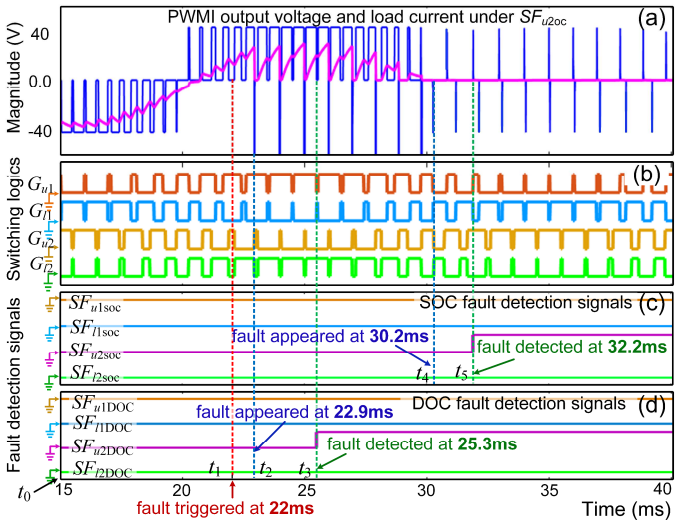


Fig. 13 Detection and identification of OC faults: PWMI output voltage with load current, switching states, and fault detection signals under SF_{u2oc} at switching frequency f_c.

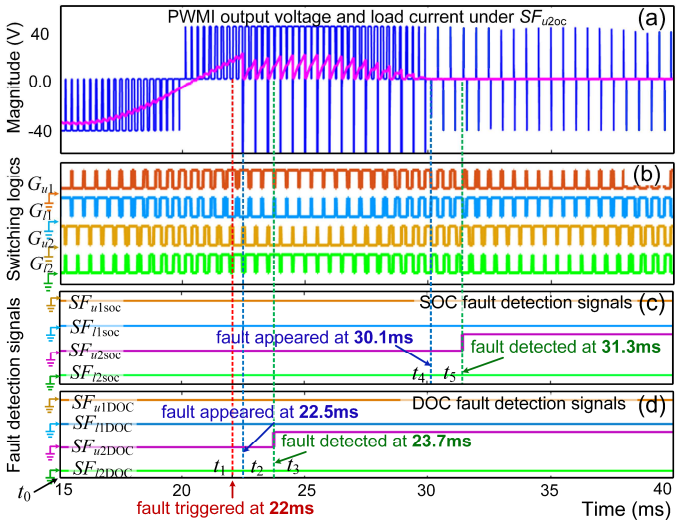


Fig. 14 Detection and identification of OC faults: PWMI output voltage with load current, switching states, and fault detection signals under SF_{u2oc} at switching frequency $2\times f_c$.

Further, the proposed detection algorithm is verified with different switching frequencies f_c, $2\times f_c$ and results are presented under open switch fault SF_{u2oc} in Figs. 13 and 14; respectively. Since the frequency of operation is increased, therefore the results plotted in Figs. 13 and 14 are captured for 25 msec (i.e., 15 – 40 msec). The open circuit fault in Sw_{u2} is triggered at t_1 = 22 msec (i.e., $t_0 - t_1$: healthy operation); the effect of diode fault occurred in PWMI output voltage at t_2 = 22.9 msec (in Fig. 13) and t_2 = 22.5 msec (in Fig. 14); thereafter DOC fault detection logic became high within couple of switching period at t_3 = 25.3 msec (in Fig. 13) and t_3 = 23.7 msec (in Fig. 14); again, at t_4 = 30.2 msec (in Fig. 13) and t_4 = 30.1 msec (in Fig. 14), the effect of switch fault occurred in PWMI output voltage and SOC fault detection logic became high within couple of switching period at t_5 = 32.2 msec (in Fig. 13) and t_5 = 31.3 msec (in Fig. 14). Therefore, for switching frequency f_c under DOC fault in Sw_{u2} is detected in 2.4 msec which is further reduced to 1.2 msec if the operating switching frequency is increased to $2\times f_c$; similarly, under SOC fault in Sw_{u2} is detected in 2.0 msec and 1.2 msec at operating switching frequencies f_c and $2\times f_c$, respectively. The fault signature in diagonally opposite switch Sw_{l1} and Sw_{u2} in an H-bridge power cell at same switching frequency are approximately identical and it can be seen from Figs. 12 and 13, respectively. It can be seen from Fig. 12 that the detection time for DOC fault in Sw_{l1} is 2.4 msec and for SOC fault in Sw_{l1} is 2.2 msec. The fault detection time is reduced with the increase of switching frequency and it is approximately twice of the switching time period. Since the fault detection algorithm is based on switching logics and PWMI output voltage therefore at higher switching frequency high speed controllers are required.

The experimental setup is shown in Fig. 15. It is developed in the laboratory to test and verify the proposed fault detection algorithm and simulated results in real time. The dSPACE 1104 digital controller is used to generate the PSPWM gating signals and implement the fault detection algorithm. The PWMI output voltage is feedback to the controller. The SOC faults are triggered manually with the help of series connected external switches, which create an interrupt to the respective gating signals. Two *legs* of the H-bridge power cell are used to test the results. For the verification of proposed algorithm more than four outputs are required, therefore two digital storage oscilloscopes (DSO's) are connected and synchronized to capture the results. The same system parameters are taken into consideration for real time implementation and verification of simulated and analytical results.

The faults in all the four switches are triggered and the resulting PWMI output voltage and load current waveform are captured as shown in Fig. 16. The counter output, diagonally opposite gating signals and fault identification signals are also presented for SF_{u1soc}, SF_{l2soc} and SF_{u2soc}, SF_{l1soc} in Figs. 16 (a), (b) and Figs. 16 (c), (d); respectively. The value of counter start increasing and reaches at defined threshold level i.e., $k_{count} = N$; whenever SOC fault is detected. As discussed in the fault identification algorithm that the counter start counts only with the non-zero values of $|v_{err}|$, i.e., the change in the actual PWMI output voltage waveshape with respect to ideal output. But from Fig. 16, it can be seen that the non-zero counter values are present during healthy operation of switches and the approximate count of the counter is $k_{count} \approx 4 = k_{healthy}$. This value is due to sensor delay, sampling delay of the controller, rise-time and fall-time switching delay of the real time system and processing delay. Therefore, to reduce the probability of false detection, the value of $N = 7$ is chosen. It can also be seen from the experimental results that diagonally opposite switch faults are approximately similar and each switch fault is detected and specific switch is identified approximately in 2.5 switching cycles.

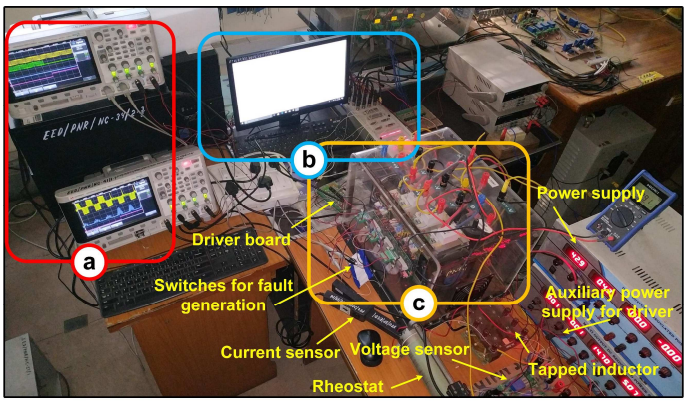


Fig. 15 Experimental setup (a) synchronized oscilloscopes, (b) computer with dSPACE 1104, (c) H-bridge cell with other system components.

0885-8993 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:44:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2021.3103851, IEEE Transactions on Power Electronics

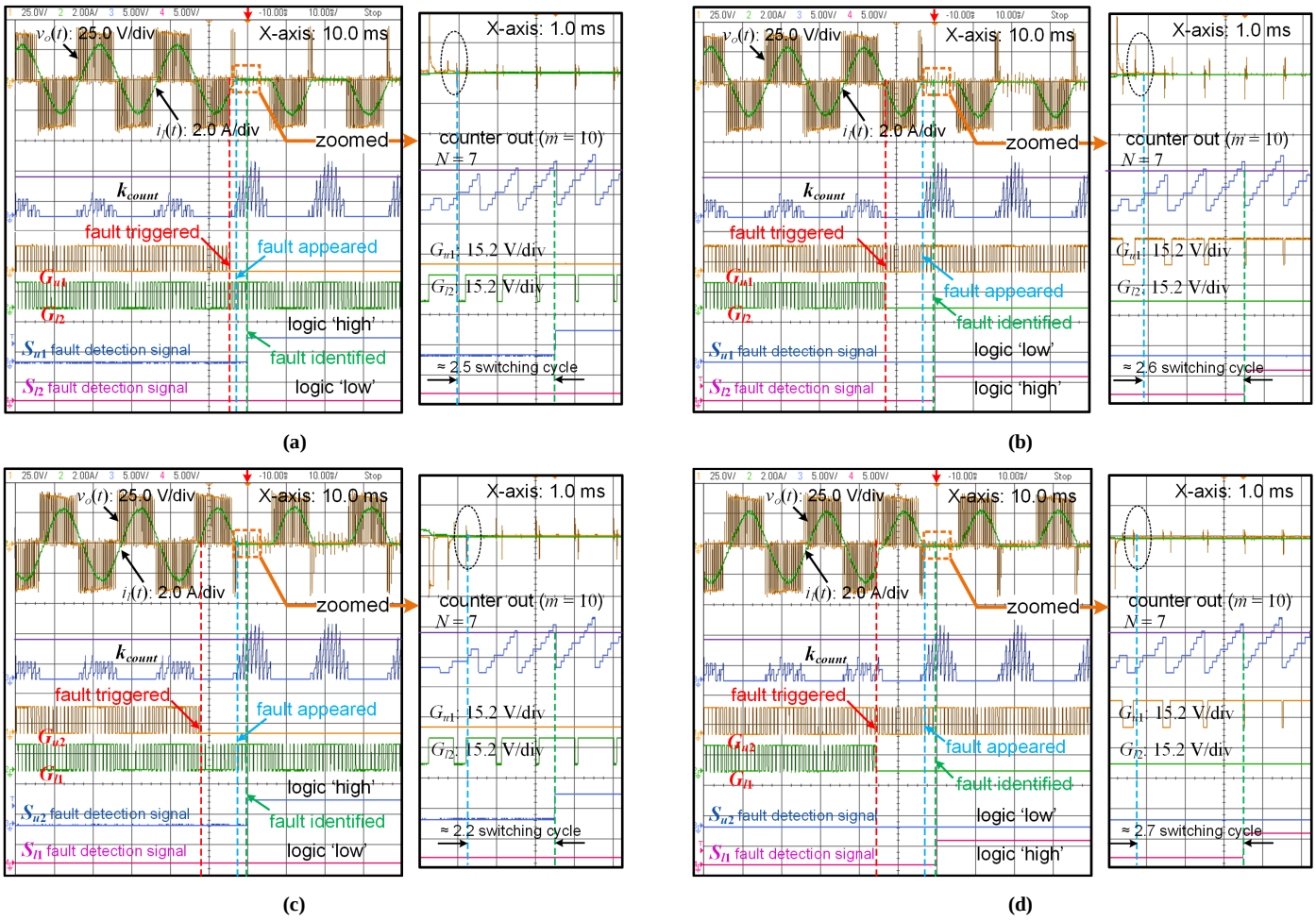


Fig. 16 PWMI output voltage, load current, counter output with threshold level $N = 7$, diagonally opposite switching logics and respective fault identification signals for (a) SF_{u1soc} with zoomed section; (b) SF_{l2soc} with zoomed section; (c) SF_{u2soc} with zoomed section; and (d) SF_{l1soc} with zoomed section.

The performance analysis of proposed switch fault identification algorithm is presented in Table III. The switch identification time period (T_{sid}) is presented for both: single switch faults and multiple switch faults. These faults are identified under several conditions for the verification of proposed algorithm, such as: open switch faults (*a*, *b*, *e*, *f*), switch-diode pair open faults (*c*, *d*), increased switching frequency of operation (*d*), different loads (*a* and *b*, *c*, *d*, *e*, *f*). It shows that the proposed fault detection and switch identification algorithm is potentially able to identify switch fault location under single switch faults, multiple switch faults, open switch faults, switch-diode pair faults and diagonaly opposite switch fault conditions. The time period between fault occurred and fault appeared at inverter output is defined as undetected region (T_d), and the value of T_d depends on the instant at which fault occurred and faulty switch location. The maximum possible value of T_d for H-bridge power cell will be $T_o/2$ (i.e., half the fundamental period). Whereas, the switch identification time period (T_{sid}) is defined as the time period between the fault appeared at inverter output and faulty swith identification. The switch identification time period (T_{sid}) for the proposed algorithm is less than $3\times T_{sw}$, where T_{sw} is switching time period. Since the magnitude of T_{sid} depends on switching time period, therefore at higher switching frequency operation results faster switch fault identification. The maximum time required for fault detection and switch fault identification using proposed algorithm is less than ($T_o/2 + 3\times T_{sw}$). The comparison between the available OCSF's detection algorithm in literature and proposed algorithm is presented in Table IV.

TABLE III
PERFORMANCE RESULT OF PROPOSED FAULT DETECTION AND SWITCH IDENTIFICATION ALGORITHM

Fault type		Description		Switch fault identification time T_{sid}	Results	Undetected region T_d / comment's
Single switch fault	*a.*	SF_{u1soc}		2.3 ms	Fig. 10 (a)	5.6 ms
		SF_{u2soc}		1.4 ms	Fig. 10 (d)	0 ms / identified under partial fault mode
	b.	SF_{u1soc}		2.0 ms	Fig. 11 (a)	0 ms
	c.	SF_{l1oc}		2.4 ms	Fig. 12	0.3 ms / switch-diode pair fault
		SF_{u2oc}		2.4 ms	Fig. 13	0.9 ms / switch-diode pair fault
	d.	SF_{u2oc}		1.2 ms	Fig. 14	0.5 ms / switch-diode pair fault and at $2\times f_{sw}$
	e.	SF_{u1soc}		2.5 ms	Fig. 16 (a)	2.0 ms
		SF_{l2soc}		2.6 ms	Fig. 16 (b)	9.0 ms
		SF_{u2soc}		2.2 ms	Fig. 16 (c)	9.5 ms
		SF_{l1soc}		2.7 ms	Fig. 16 (d)	10.0 ms
Multiple switch fault	*f.*	*leg*-1	SF_{u1soc}	2.1 ms	Fig. 11 (d)	0 ms
			SF_{l1soc}	2.35 ms		8.05 ms

0885-8993 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:44:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2021.3103851, IEEE Transactions on Power Electronics

TABLE IV
COMPARISON BETWEEN PROPOSED FAULT DETECTION AND SWITCH IDENTIFICATION ALGORITHM WITH AVAILABLE DETECTION METHODS IN LITERATURE

Algorithms from literature	(T_d+T_{sid}) for single switch fault	(T_d+T_{sid}) for multiple switch fault	(T_d+T_{sid}) for switch-diode pair / diode fault	DOSFI under similar PWMI output voltage	Detection criteria	Computational complexity
[35]	~6×T_o	Not possible	Not possible	Not possible	FFT and principle component analysis (PCA)	High
[21]	~6×T_o	Not possible	Not possible	Not possible	Output phase voltage and PCA	High
[19]	~3×T_o	Not possible	Not possible	Not possible	Slope calculations and comparison	Moderate
[24]	~5×T_o	Not possible	Not possible	Not possible	Residual calculations and comparison	Moderate
[22]	~1×T_o	~1×T_o	Not possible	Not possible	Reference current predition and comparison	Moderate
[13]	~1×T_o	~1×T_o	Not possible	Not possible	Comparison of two variables	Low
Proposed method	~{(T_o/2)+(3×T_{sw})}	~{(T_o/2)+(3×T_{sw})}	~{(T_o/2)+(3×T_{sw})}	Possible	Comparison of variables	Low

T_{sid} = switch identification period; T_d = undetected region; T_o = fundamental period; T_{sw} = switching period; # DOSFI = diagonally opposite switch fault identification.

VI. CONCLUSION

The semiconductor OCSF's are less critical than short circuit faults for power converter operations but the severity increases with type of OCSF's. The output qualities and productivity of the power electronics systems are affected and it may also damage the complete system due to large voltage spikes during the switching transitions. Therefore, a fast OCSF detection algorithm is required to identify the switch faults within minimum switching transitions.

In this paper, a fast OCSF's detection logic design is proposed for the H-bridge power cell. The OCSF can be detected and faulty switch can be identify within (T_o/2 + 3×T_{sw}) period for H-bridge cell. In H-bridge power cell, the fault signatures of diagonally opposite switch faults are approximately similar therefore it is difficult to identify unique switch fault location. This paper presents a detection algorithm with unique identification for every faulty switch within H-bridge power cell. The non-ideal characteristics of semiconductor switches are used for the identification of OCSF's. The method compares the actual and ideal PWMI output voltage, the instant of voltage spikes due to non-ideal switching characteristics with respect to switching states. The proposed algorithm can be applied for the identification of both single open switch fault and multiple open switch faults. It can also be applied for identification of open switch faults and switch-diode pair fault. The proposed technique requires only one voltage sensor per H-bridge power cell for SOC fault identification whereas it requires one voltage and one current sensor per H-bridge power cell for all types of OCSF's. The PSPWM technique is used for three-level voltage output generation but the proposed method can be further implemented for other modulation techniques and the design can also be extended for the cascaded H-bridge based multilevel inverters.

REFERENCES

[1] S. Yang, A. Bryant, P. Mawby, D. Xiang, L. Ran and P. Tavner, "An industry-based survey of reliability in power electronic converters," *IEEE Trans. Ind. Appl.*, vol. 47, no. 3, pp. 1441-1451, May-June 2011.

[2] U. Choi, F. Blaabjerg and K. Lee, "Study and handling methods of power IGBT module failures in power electronic converter systems," *IEEE Trans. Power Electron.*, vol. 30, no. 5, pp. 2517-2533, May 2015.

[3] J. Rodriguez, S. Bernet, B. Wu, J. O. Pontt and S. Kouro, "Multilevel voltage-source-converter topologies for industrial medium-voltage drives," *IEEE Trans. Ind. Electron.*, vol. 54, no. 6, pp. 2930-2945, Dec. 2007.

[4] S. Daher, J. Schmid and F. L. M. Antunes, "Multilevel inverter topologies for stand-alone PV systems," *IEEE Trans. Ind. Electron.*, vol. 55, no. 7, pp. 2703-2712, July 2008.

[5] J. Rodriguez, Jih-Sheng Lai and Fang Zheng Peng, "Multilevel inverters: a survey of topologies, controls, and applications," *IEEE Trans. Ind. Electron.*, vol. 49, no. 4, pp. 724-738, Aug. 2002.

[6] B. Lu and S. K. Sharma, "A literature review of IGBT fault diagnostic and protection methods for power inverters," *IEEE Trans. Ind. Appl.*, vol. 45, no. 5, pp. 1770-1777, Sept.-Oct. 2009.

[7] B. Mirafzal, "Survey of fault-tolerance techniques for three-phase voltage source inverters," *IEEE Trans. Ind. Electron.*, vol. 61, no. 10, pp. 5192-5202, Oct. 2014.

[8] W. Zhang, D. Xu, P. N. Enjeti, H. Li, J. T. Hawke and H. S. Krishnamoorthy, "Survey on fault-tolerant techniques for power electronic converters," *IEEE Trans. Power Electron.*, vol. 29, no. 12, pp. 6319-6331, Dec. 2014.

[9] A. A. Stonier and B. Lehman, "An intelligent-based fault-tolerant system for solar-fed cascaded multilevel inverters," *IEEE Trans. Energy Conv.*, vol. 33, no. 3, pp. 1047-1057, Sept. 2018.

[10] H. K. Jahan, F. Panahandeh, M. Abapour and S. Tohidi, "Reconfigurable multilevel inverter with fault-tolerant ability," *IEEE Trans. Power Electron.*, vol. 33, no. 9, pp. 7880-7893, Sept. 2018.

[11] J. Wang and Y. Tang, "A fault-tolerant operation method for medium voltage modular multilevel converters with phase-shifted carrier modulation," *IEEE Trans. Power Electron.*, vol. 34, no. 10, pp. 9459-9470, Oct. 2019.

[12] R. S. Chokhawala, J. Catt and L. Kiraly, "A discussion on IGBT short-circuit behavior and fault protection schemes," *IEEE Trans. Ind. Appl.*, vol. 31, no. 2, pp. 256-263, March-April 1995.

[13] N. B. Y. Gorla, S. Kolluri, M. Chai and S. K. Panda, "A novel open-circuit fault detection and localization scheme for cascaded H-bridge stage of a three-stage solid-state transformer," *IEEE Trans. Power Electron.*, vol. 36, no. 8, pp. 8713-8729, Aug. 2021.

[14] G. N. Brahmendra Yadav and S. K. Panda, "Survey of open-circuit fault detection and localization methods applicable to cascaded H-bridge multilevel converters," The 46th Annual Conference of the IEEE Industrial Electronics Society (IECON 2020), Singapore, pp. 3988-3994, 2020.

[15] D. Sadik et al., "Short-circuit protection circuits for silicon-carbide power transistors," *IEEE Trans. Ind. Electron.*, vol. 63, no. 4, pp. 1995-2004, April 2016.

[16] M. A. Rodríguez-Blanco et al., "A failure-detection strategy for IGBT based on gate-voltage behavior applied to a motor drive system," *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 1625-1633, May 2011.

[17] Z. Wang, X. Shi, L. M. Tolbert, F. F. Wang, and B. J. Blalock, "A di/dt feedback-based active gate driver for smart switching and fast overcurrent protection of igbt modules," *IEEE Trans. Power Electron.*, vol. 29, no. 7, pp. 3720–3732, 2014.

[18] S. Haghnazari, M. Khodabandeh, and M. R. Zolghadri, "Fast fault detection method for modular multilevel converter semiconductor power switches," *IET Power Electron.*, vol. 9, no. 2, pp. 165–174, 2016.

[19] H. Sim, J. Lee and K. Lee, "Detecting open-switch faults: using asymmetric zero-voltage switching states," *IEEE Ind. Appl. Mag.*, vol. 22, no. 2, pp. 27-37, Mar.-Apr. 2016.

[20] J. Lamb and B. Mirafzal, "Open-circuit IGBT fault detection and location isolation for cascaded multilevel converters," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 4846-4856, June 2017.

[21] S. Khomfoi and L. M. Tolbert, "Fault diagnosis and reconfiguration for multilevel inverter drive using AI-based techniques," *IEEE Trans. Ind. Electron.*, vol. 54, no. 6, pp. 2954-2968, Dec. 2007.

[22] M. Chai, N. B. Y. Gorla and S. K. Panda, "Fault detection and localization for cascaded H-bridge multilevel converter with model predictive control," *IEEE Trans. Power Electron.*, vol. 35, no. 10, pp. 10109-10120, Oct. 2020.

0885-8993 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:44:19 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2021.3103851, IEEE Transactions on Power Electronics

[23] D. Zhou, S. Yang and Y. Tang, "A voltage-based open-circuit fault detection and isolation approach for modular multilevel converters with model-predictive control," *IEEE Trans. Power Electron.*, vol. 33, no. 11, pp. 9866-9874, Nov. 2018.

[24] K. Thantirige, S. Mukherjee, M. A. Zagrodnik, C. Gajanayake, A. K. Gupta and S. K. Panda, "Reliable detection of open-circuit faults in cascaded H-bridge multilevel inverter via current residual analysis," IEEE Transportation Electrification Conference (ITEC-India), Pune, India, pp. 1-6, 2017.

[25] P. Lezana, R. Aguilera, and J. Rodr´ıguez, "Fault detection on multicell converter based on output voltage frequency analysis," *IEEE Trans. Ind. Electron.*, vol. 56, no. 6, pp. 2275–2283, 2009.

[26] I. Abari, A. Lahouar, M. Hamouda, J. B. H. Slama, and K. Al-Haddad, "Fault detection methods for three-level npc inverter based on dc-bus electromagnetic signatures," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5224–5236, 2017.

[27] Y. Chen, X. Pei, S. Nie, and Y. Kang, "Monitoring and diagnosis for the dc–dc converter using the magnetic near field waveform," *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 1634–1647, 2010.

[28] T. Jing and Z. Chang-Hao, "The model of IGBT as a power switch," 2011 International Conference on Electronics, Communications and Control (ICECC), Ningbo, pp. 1720-1723, 2011.

[29] M. A. Rodríguez-Blanco et al., "A failure-detection strategy for IGBT based on gate-voltage behavior applied to a motor drive system," *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 1625-1633, May 2011.

[30] P. Xue, G. Fu and D. Zhang, "Modeling inductive switching characteristics of high-speed buffer layer IGBT," *IEEE Trans. Power Electron.*, vol. 32, no. 4, pp. 3075-3087, April 2017.

[31] M. Kumar and R. Gupta, "Time-domain analysis of sampling effect in DPWM of DC–DC converters," *IEEE Trans. Ind. Electron.*, vol. 62, no. 11, pp. 6915–6924, Nov. 2015.

[32] M. Kumar and R. Gupta, "Sampling effect characterization of digital SPWM of VSI in time domain," *IEEE Trans. Ind. Electron.*, vol. 63, no. 7, pp. 4150–4159, Jul. 2016.

[33] D. G. Holmes and T. A. Lipo, Pulse Width Modulation for Power Converters. Piscataway, NJ, USA: *IEEE Press*, 2003.

[34] M. Kumar and R. Gupta, "Time-domain characterisation of multicarrierbased digital SPWM of multilevel VSI," *IET Power Electron.*, vol. 11, no. 1, pp. 100–109, Jan. 2018.

[35] T. Wang, H. Xu, J. Han, E. Elbouchikhi and M. E. H. Benbouzid, "Cascaded H-bridge multilevel inverter system fault diagnosis using a PCA and multiclass relevance vector machine approach," *IEEE Trans. Power Electron.*, vol. 30, no. 12, pp. 7006-7018, Dec. 2015.



Mayank Kumar (Senior Member IEEE) was born in Jaunpur, India, in 1986. He received the B.Tech. degree in electronics and communication engineering from Uttar Pradesh Technical University, Lucknow, India, in 2010, the M.Tech. degree and the Ph.D. degree in electrical engineering from Motilal Nehru National Institute of Technology, Allahabad, India, in 2013 and 2017, respectively. He is currently an Assistant Professor with the Department of Electrical Engineering, Delhi Technological University, Delhi, India.

His research interests include digital control of power electronic converters, FPGAs, switching techniques of dc–dc, dc–ac, and ac–ac converters, modelling and control of switched power electronics circuits, fault tolerant converters, solar power conversion, etc.

0885-8993 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:44:19 UTC from IEEE Xplore. Restrictions apply.

Tribological characterization of pultruded hybrid glass-jute fibre reinforced plastic composites from room temperature to 75°C

Navin Kumar
Department of Mechanical Engineering, National Institute of Technology Kurukshetra, Kurukshetra, India and
Department of Mechanical Engineering, Greater Noida Institute of Technology, Greater Noida, India
R.S. Walia
Production and Industrial Engineering Department, Punjab Engineering College, Chandigarh, India, and
Surjit Angra
Department of Mechanical Engineering, National Institute of Technology, Kurukshetra, India

Abstract

Purpose – The aim of the present study is to find the tribological properties of newly developed polyester-based hybrid glass-jute fibre reinforced plastic composites loaded with different weight per cent of hybrid filler particles were investigated under a dry sliding medium from room temperature to 75°C.

Design/methodology/approach – The study was carried out using a pin-on-disc wear test set-up. The design of experiments was carried out in a controlled way using a central composite design based on response surface methodology to observe the effect of various parameters i.e. sliding velocity, sliding distance, the temperature of counterface and different applied load conditions during dry-sliding.

Findings – The maximum wear resistance was found at 9 Wt% loading of filler, 4 ms^{-1} sliding velocity, 30 N applied load, 54°C temperature of the counterface and 1,100 m sliding distance condition. Optimum values of hybrid filler loading, sliding velocity, applied load, the temperature of the counterface and sliding distance for the minimum coefficient of friction value and minimum friction force are 9 Wt%, 4 ms^{-1}, 30 N, 54° C, 1,100 m and 12 Wt%, 3 ms^{-1}, 20 N, 59°C and 1,100 m, respectively. The worn surface morphology was studied using scanning electron microscope, for wear dominant mechanisms.

Originality/value – The tribological properties of newly developed polyester-based hybrid glass-jute fibre reinforced plastic composites loaded with different weight % of hybrid filler particles, were investigated under dry sliding medium from room temperature to 75°C has not been attempted yet.

Keywords Wear, Hybrid GJFRP, Hybrid filler, Pultrusion, Coefficient of friction, Response surface methodology (RSM), Dry sliding

Paper type Research paper

Abbreviation

COF Values	= Coefficient of friction values;
GJFRP	= Glass-jute fibres reinforced plastics;
HF 3, 6, 9, 12, 15	= Hybrid composite with 3, 6, 9, 12, 15 Wt% filler loading;
ANOVA	= Analysis of variance;
Wear 1, wear 2. . ., wear 32	= Wear loss during run1. . ..32;
COF 1, COF 2,. . ., COF 32	= Coefficient of friction values for run1. . .32;
FRP	= Fibre reinforced polymer;
GFRP	= Glass Fibre reinforced polymer;
Wt%	= Weight%;
HRC	= Hardness Rockwell C; and
Ra	= Roughness average of surface.

1. Introduction

In recent years, hybrid fibre reinforced polymer (FRP) composites have been developed using hybrid fibres and hybrid fillers. Hybridization of fibres is an effort to replace the synthetic fibres with natural fibres partially or completely due to environment-related concerns and their safe and non-harmful disposal at the end of life. Nowadays, FRP composites made equipment and component parts are gaining popularity in marine, aerospace, automobile and other structural applications, but major concern related to their application is of

The current issue and full text archive of this journal is available on Emerald Insight at: **https://www.emerald.com/insight/1708-5284.htm**



World Journal of Engineering
© Emerald Publishing Limited [ISSN 1708-5284]
[DOI 10.1108/WJE-03-2021-0147]

The support of resources of Precision Manufacturing Lab, Delhi Technological University, Delhi, India is gratefully acknowledged.

Received 14 March 2021
Revised 15 May 2021
Accepted 7 June 2021

surface degradation. Abrasive surroundings, different sliding and rolling media, cyclic loading conditions and temperature-based applications may initiate the surface degradation of FRP made equipment and components. These factors may further start cracks formation on the matrix and may reduce the material life. FRP composite materials should not only be cost-effective, light in weight and easy to process, it should also have a low coefficient of friction, high wear resistance and increased life span for its use in relative sliding motion condition (Joshi *et al.*, 2004; Karthikeyan *et al.*, 2017; Koronis *et al.*, 2013). Under loading conditions and temperature-based applications attracted the researchers for wear resistance analysis of hybrid FRP composites. It was found that specific wear rate in acrylic-based wood fibre composites was lowest with 50 Wt% of fibre loading (Akpan *et al.*, 2018). During dry sliding, the wear resistance was increased and the coefficient of friction was decreased due to the 10 Wt%–20 Wt% amount of filler fly-ash addition in glass FRP composites (Chauhan *et al.*, 2010). Coefficient of friction values decreased under dry sliding and oil-lubricated sliding conditions, while maximum values were during sliding in inert gas (argon) medium for glass fibre reinforced epoxy resin (Agrawal *et al.*, 2016). A significant decrease in friction and wear values were observed in carbon fibre reinforced epoxy and polyethersulfone and glass fibre reinforced polypropylene composites when contacted with a diamond-like carbon layer (ta-C) coated steel (Andrich *et al.*, 2013). Friction coefficient value was increased in the carbon fibre reinforced composites due to the increase of normal load and sliding velocity (Kaya and Parlar, 2020). During the tribological behaviour study of epoxy based hybrid composites reinforced with flax, jute and banana fibres, it was observed that wear behaviour of hybrid FRP composites is greatly affected by the hardness, volume fraction and speed. Speed and hardness have positive effect while volume fraction negatively contributes on the wear behaviour of hybrid FRP composites (Mangalaraja *et al.*, 2019; Babu *et al.*, 2019). Yallew *et al.* (2014) observed 65% reduction in specific wear rate and 3.5%–45% decrease in coefficient of friction during the tribological behaviour study of jute fabric reinforced polypropylene composites. The outcomes were possible due to the loading of woven jute fabric in to the polypropylene matrix. Latha *et al.* (2016) concluded that tribological, as well as mechanical properties of the hybrid composites are greatly affected by the layer sequence in hand lay-up technique made laminates. Glass fibre loading in bamboo composites enhanced its overall properties significantly.

The wear rate of carbon FRP hybrid nano-composites was reduced and was minimum as 2.22 mg/Nm when 3 Wt% of each nanoclay and nano-ZnO were loaded. The study results suggested the addition of nano-filler particles in the FRP composites to enhance the wear resistance of such materials (Tripathi and Ambekar, 2019). The tribological behaviour of glass fibre reinforced epoxy composites filled with red brick dust was conducted and concluded that normal load and sliding velocity were the two most prominent factors which greatly affected the coefficient of friction and wear rate (Pati and Satpathy, 2021) Glass-jute reinforced hybrid composites loaded with cenosphere filler showed maximum erosion at (45°–60°) impingement angle and impact velocity was increasing the erosion rate (Dalbehera and Acharya, 2015). While studying the epoxy based hybrid composites wear behaviour, it was observed that reinforcement of both short human hair fibres and solid glass microspheres improved the wear performance of the material (Nanda and Satapathy, 2020). The loading of specific fillers and fibres in a particular amount, sliding in a particular direction and interaction of wear sample with a counterface of some specific material reduced the coefficient of friction and wear values of polymer based composites (Giltrow and Lancaster, 1970; Ruggiero *et al.*, 2015; Chairman and Kumaresh Babu, 2013; Monfared *et al.*, 2018). In automobile industry, FRPs are gaining popularity due to their low coefficient of friction and very less sliding wear. Polymeric composites are found with applications in clutches, gears, bush bearing and wheels. Hence, knowledge of their tribological behaviour is must before their application.

Although, significant research work has been carried out for the advancement of FRP composites, yet development of hybrid glass-jute fibres reinforced plastic composites using pultrusion process and loading biodegradable hybrid fillers remain uncharted. The replacement of synthetic fibres (glass fibres) with maximum use of natural fibres (jute fibres) and loading of carbonized form of eggshells, coconut coir and carbon black as a hybrid filler material made the developed hybrid composites an eco-friendly one. The maximum use of natural fibres and bio-hazardous waste material eggshells in combination with agricultural waste coconut coir in the development of such advanced materials not only lowers down the cost of the material but also removes the pollution related concerns. The present study not only shows the development of eco-friendly hybrid composite materials but also explores their temperature based tribological behaviour study, which makes it a novel work.

2. Experimental

2.1 Material preparation

A mixture of powder consisting of 6 Wt% of carbonized eggshell ash, 6 Wt% of coconut coir ash and 9 Wt% of carbon black, which is individually the weight per cent of the total weight of resin used in a single development run, was taken as hybrid filler in 75 micron particle size. Other materials used are hybrid fibre, catalyst and matrix material with fix amount as shown in Table 1.

Six different compositions of composites were developed using materials as per Table 1 and loading fillers amounts as (0, 3, 6, 9, 12, 15) Wt%. Three samples of each type composites with a fixed amount of fillers loading were developed. Total developed 18 composite specimens were tested for tensile

Table 1 Constant parameters and materials amount for experiment

Hot cylindrical die temperature	180°C
Speed of pultrusion of steel strip	60 mm min^{-1}
(Glass-Jute) hybrid fibre in untreated form	300 g (225 g jute fibre + 75 g glass fibre)
Polyester resin weight	750 g [weight of (glass-jute) hybrid fibre) × 2.5]
Cobalt naphthelate	3.75 g (0.5% of resin weight)
Catalyst (Methyl ethyl ketone peroxide)	7.5 g (1% of resin weight)

strength and average results were selected for discussion. The obtained tensile strength results of GJFRP composites at (0, 3, 6, 9, 12, 15) Wt% fillers loading are (60.33, 66.0, 72.98, 88.37, 73.14, 65.33) MPa, respectively. The observed tensile strength was maximum at 9 Wt% fillers loading, which is confirmed by previous study results (Kumar *et al.*, 2021). Hence, 9 Wt% fillers loading can be considered as an ideal value for fillers addition.

2.2 Specimen preparation and test details

Hybrid GJFRP composites were developed using pultrusion process, which is found as most suitable manufacturing process by many researchers Gupta *et al.*, 2014, 2015, 2016, 2019a, 2019b). The used pultrusion set-up is shown in Figure 1. The main components of pultrusion set-up are hot cylindrical shape die, pre-former and puller assembly. Hot cylindrical shape die is connected with a temperature controller. The puller assembly consists of three-phase A.C. motor (1 HP), three-sets of pulling rollers and speed controlling arrangement. During the development of hybrid GJFRP composites, preformed bundles of glass and jute fibres of 1,500 mm length were given a bath of hybrid filler blended polyester resin. These resin wetted bundles of hybrid fibres were then pulled by pulling assembly in the pultrusion set-up against the hot die at 180°C with a constant speed of 60 mm.min^{-1} passing firstly through cold die or performer, getting a suitable shape and removing the extra amount of resin. Cobalt naphthenate of 6% concentration and MEKP were used as accelerator and catalyst in 0.5 Wt% and 1 Wt% amount of resin weight, respectively, in each

Figure 1 Pultrusion set-up

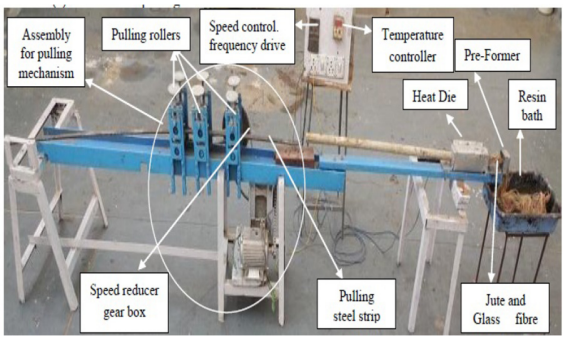


Figure 2 High temperature based rotary tribometer (DUCOM, India)

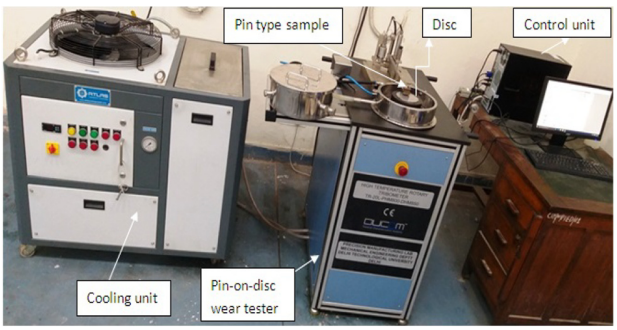


Figure 3 Wear sample (pin)



experiment. Five different compositions of composites were developed with fillers amounts as (3, 6, 9, 12, 15) Wt%.

The wear and friction behaviour study of these developed composites was conducted as per ASTM: G99-05 using a pin-on-disc tribometer as shown in Figure 2. The pin type wear test specimens were prepared as shown in Figure 3 with 12 mm diameter and 45 mm length. An EN-31 made steel disc of 8 mm thickness, 100 mm diameter, 62HRc hardness and with a smooth and even surface with surface roughness as Ra = 0.7 was used as a counterface as shown in Figures 4 and 5.

Five process parameters were selected for the wear experiments as load (N), sliding velocity (m/s), sliding distance (m), hybrid filler wt (per cent) and temperature (°C). Track diameters on steel disc were kept constant as 60 mm for all the experiments. To achieve the best contact between the surface

Figure 4 Disc (EN 31)

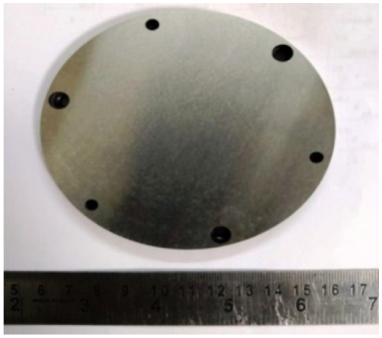


Figure 5 Wear sample position on the disc

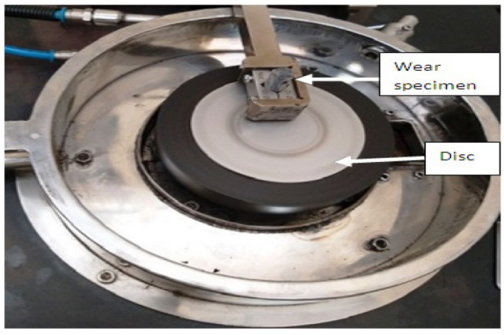


Table 2 Central composite design and observed experimental results

Std.	Run	Hybrid filler (%)	Sliding velocity (m/s)	Load (N)	Temperature (°C)	Sliding distance (m)	Coefficient of friction (Fraction)	Wear (Micron)	Friction Force (N)
1	30	6	3	20	45	1,400	0.29	92.84	18.79
2	9	12	3	20	45	800	0.18	12.35	3.57
3	32	6	5	20	45	800	0.24	33.15	7.55
4	19	12	5	20	45	1,400	0.32	25.43	9.67
5	15	6	3	40	45	800	0.28	20.00	11.67
6	29	12	3	40	45	1,400	0.47	122.92	15.36
7	22	6	5	40	45	1,400	0.53	11.96	5.67
8	31	12	5	40	45	800	0.26	79.94	12.17
9	2	6	3	20	65	800	0.13	28.56	9.14
10	8	12	3	20	65	1,400	0.21	21.33	10.25
11	18	6	5	20	65	1,400	0.33	24.46	3.33
12	10	12	5	20	65	800	0.20	58.67	8.11
13	3	6	3	40	65	1,400	0.47	51.12	16.94
14	7	12	3	40	65	800	0.39	33.23	10.12
15	1	6	5	40	65	800	0.39	61.35	6.67
16	27	12	5	40	65	1,400	0.30	36.00	12.67
17	17	3	4	30	55	1,100	0.40	80.67	10.90
18	25	15	4	30	55	1,100	0.35	99.77	12.34
19	11	9	2	30	55	1,100	0.30	59.34	14.44
20	5	9	6	30	55	1,100	0.34	36.34	8.51
21	26	9	4	10	55	1,100	0.10	6.67	4.33
22	23	9	4	50	55	1,100	0.40	38.67	9.47
23	28	9	4	30	35	1,100	0.25	13.22	9.42
24	13	9	4	30	75	1,100	0.16	9.65	7.67
25	4	9	4	30	55	500	0.18	11.23	5.47
26	12	9	4	30	55	1,700	0.38	25.12	12.56
27	14	9	4	30	55	1,100	0.24	18.77	9.66
28	24	9	4	30	55	1,100	0.24	11.34	5.18
29	6	9	4	30	55	1,100	0.11	8.00	4.67
30	16	9	4	30	55	1,100	0.24	8.67	4.77
31	20	9	4	30	55	1,100	0.23	9.45	5.89
32	21	9	4	30	55	1,100	0.23	12.45	5.56

of the disc and test sample, both were cleaned with acetone soaked soft paper and dried well before proceeding for each test run. Experiments were carried out in dry sliding medium as per central composite design. The responses for friction force, wear and coefficient of friction were observed as shown in Table 2 below.

3. Results and discussion

The ANOVA tables for coefficient of friction (COF) values, wear and frictional force are shown in Tables 3, 4 and 5. The range of desirability factor is 0 to 1, in which the highest value shows a highly desirable factor and vice versa. The optimum variable setting was found when the process variables were with ultimate desirability. Insignificant interactions were not included in the models for COF values, wear and friction force and only the main factors with their mutual interactions were included.

3.1 Coefficient of friction

The contour and three-dimensional response surface plots were the key factors in analyzing the effects of variable parameters on the coefficient of friction as an outcome. These plots for the effect of variable parameters on COF value were established as per the fitted model. Table 2 shows the central composite design for the actual factors of the experiments and also the results of various experiments for COF values along with actual factors.

The analysis of variance table for COF values shows the model as significant as F-value of 19.65. p-values less than 0.05 showed significant terms of the model. The possibility of getting the larger F-values was 0.01% and possible responsible factor was found to be noise. From the model terms such as A, B, C, D, E, AB, BC, DE, A^2, B^2 and E^2 were found significant with their percentage contribution as 2.23%, 0.81%, 39.68%, 1.36%, 20.01%, 4.28%, 2.83%, 3.98%, 15.28%, 7.17% and 2.44%, respectively. Values with more than 0.1 gave a clue about the insignificant nature of the model terms. If a large number of such insignificant terms are present, then the method of model reduction can be implemented to improve the model. The results show the F-value as 0.39, which represented the non-significant nature of lack of fit. In that case, there was

Table 3 Pooled ANOVA table for COF

Source	C.V	A.V	S.S	D.O.F	M.S	*F*- value	Probability > *F*
Model	+0.2208	−0.290478	0.3233	11	0.0294	19.65	<0.0001 (Significant)
A	−0.0173	−0.048663	0.0072	1	0.0072	4.79	0.0406
B	+0.0106	−0.053771	0.0027	1	0.0027	1.81	0.1942
C	+0.0741	+0.017090	0.1317	1	0.1317	88.00	<0.0001
D	−0.0137	+0.009174	0.0045	1	0.0045	3.02	0.0976
E	+0.0526	+0.000298	0.0664	1	0.0664	44.37	<0.0001
AB	−0.0298	−0.009924	0.0142	1	0.0142	9.48	0.0059
BC	−0.0242	−0.002421	0.0094	1	0.0094	6.27	0.0211
DE	−0.0288	−9.58721	0.0132	1	0.0132	8.85	0.0075
A^2	+0.0413	+0.004589	0.0507	1	0.0507	33.88	<0.0001
B^2	+0.0283	+0.028290	0.0238	1	0.0238	15.90	0.0007
E^2	+0.0165	+1.83772	0.0081	1	0.0081	5.43	0.0303
Residual			0.0299	20	0.0015		
Lack of fit			0.0161	15	0.0011	0.3907	0.9275 (non-significant)
Pure Error			0.0138	5	0.0028		
Cor Total			0.3533	31	C.V- coded value		
Std. Dev.	0.04		R^2	0.92	A.V- actual value		
Mean	0.29		Adjusted R^2	0.87	S.S- sum of squares		
C.V. %	13.55		Predicted R^2	0.82	D.O.F- degree of freedom		
Press	0.0652		Adeq precision	18.44	M.S- mean square		

Table 4 Pooled ANOVA table for wear

Source	C.V	A.V	S.S	D.O.F	M.S	*F*- value	Probability > *F*
Model	+12.98	+100.29700	28752.04	14	2053.72	90.83	<0.0001 (Significant)
A	+4.36	−44.08338	456.17	1	456.17	20.18	0.0003
B	−4.06	−63.07147	395.21	1	395.21	17.48	0.0006
C	+7.66	−4.36884	1406.60	1	1406.60	62.21	< 0.0001
D	−3.79	+1.46833	345.02	1	345.02	15.26	0.0011
E	+3.61	+0.465919	312.46	1	312.46	13.82	0.0017
AB	+4.49	+1.49606	322.30	1	322.30	14.25	0.0015
AC	+11.81	+0.393544	2230.23	1	2230.23	98.64	< 0.0001
AD	−6.18	−0.206119	611.78	1	611.78	27.06	< 0.0001
BD	+8.99	+0.899064	1293.31	1	1293.31	57.20	< 0.0001
BE	−20.58	−0.068614	6779.43	1	6779.43	299.84	< 0.0001
DE	−9.79	−0.003262	1532.65	1	1532.65	67.79	< 0.0001
A^2	+19.54	+2.17125	11346.75	1	11346.75	501.84	< 0.0001
B^2	+8.95	+8.94703	2378.61	1	2378.61	105.20	< 0.0001
C^2	+2.65	+0.026542	209.33	1	209.33	9.26	0.0074
Residual			384.38	17	22.61		
Lack of fit			306.18	12	25.52	1.63	0.3076 (non-significant)
Pure error			78.20	5	15.64		
Cor total			29136.41	31	C.V- coded value		
Std. dev.	4.76		R^2	0.99	A.V- actual value		
Mean	36.33		Adjusted R^2	0.98	S.S- sum of squares		
C.V. %	13.09		Predicted R^2	0.95	D.O.F- degree of freedom		
Press	1457.40		Adeq precision	35.34	M.S- mean square		

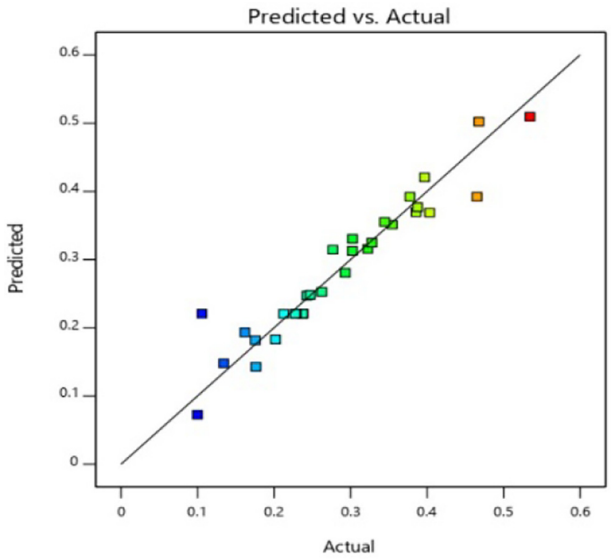
Table 5 Pooled ANOVA table for friction force

Source	C.V	A.V	S.S	D.O.F	M.S	*F*- value	Probability > *F*
Model	+6.19	+74.53581	444.63	12	37.05	19.58	<0.0001
A	+0.2106	−6.71302	1.06	1	1.06	0.5626	0.4624
B	−1.74	−12.21486	72.98	1	72.98	38.56	<0.0001
C	+1.30	−0.181400	40.38	1	40.38	21.34	0.0002
D	−0.4460	−0.680507	4.77	1	4.77	2.52	0.1288
E	+1.58	+0.013249	59.78	1	59.78	31.59	<0.0001
AB	+2.29	+0.762608	83.75	1	83.75	44.25	<0.0001
AC	+1.04	+0.034568	17.21	1	17.21	9.09	0.0071
BE	−1.88	−0.006253	56.31	1	56.31	29.75	<0.0001
A^2	+1.35	+0.149764	53.66	1	53.66	28.36	<0.0001
B^2	+1.31	+1.31075	50.75	1	50.75	26.82	<0.0001
D^2	+0.5781	+0.005781	9.87	1	9.87	5.22	0.0341
E^2	+0.6965	+7.73835E-06	14.33	1	14.33	7.57	0.0127
Residual			35.96	19	1.89		
Lack of fit			18.35	14	1.31	0.3724	0.9336
Pure error			17.60	5	3.52		
Cor total			480.58	31	C.V- coded value		
Std. dev.	1.38		R^2	0.93	A.V- actual value		
Mean	9.14		Adjusted R^2	0.88	S.S- sum of squares		
C.V. %	15.05		Predicted R^2	0.83	D.O.F- degree of freedom		
Press	81.91		Adeq Precision	15.95	M.S- mean square		

92.75% possibility for a larger value of lack of fit, which might be due to noise. The predicted R^2 of 0.82 represented a logical relationship with an adjusted R^2 of 0.87. This indicated the difference of less than 0.2. Value of adeq precision was used to find the signal to noise ratio. If the ratio exceeded 4, then it was the most suitable condition. The present model has a ratio of 18.44 directed towards the best suitability signal for the analysis.

Figure 6 indicates the superb fitness of the regression model. The outcomes were found to be equivalent to the predicted values and actual V/S predicted values of COF plot can be taken into consideration. The observed value points were found in close correlation with line of prediction indicating that all the outcomes of experimental work were in agreement with the predicted results.

Figure 6 Plot of actual V/S predicted value of COF



The coded equation for the coefficient of friction model is as shown below:

$$\begin{aligned}\text{Coefficient of friction} = {} & 0.220775 - 0.0172838 * A + 0.0106077 \\ & * B + 0.0740668 * C - 0.0137219 * D \\ & + 0.0525935 * E - 0.0297723 * AB \\ & - 0.024209 * BC - 0.0287616 * DE \\ & + 0.041299 * A^2 + 0.0282903 * B^2 \\ & + 0.0165395 * E^2 \end{aligned} \quad (1)$$

Figure 3 indicates that actual values were found in agreement with predicted values for minimization of COF values with a very little deviation. Various dry sliding wear test parameters showed their integrated effect on the COF values for hybrid GJFR plastic composites as shown in three-dimensional response surface plots (Figures 7–9). The joint effect of hybrid filler per cent and sliding velocity on COF can be elaborated with the help of Figure 7, which indicates that increase in hybrid filler per cent, increased COF initially and then decreased as these filler particles behaved like a solid lubricant. The presence of filler particles between the contacting surfaces was more with the increased filler per cent (Suresha, 2006). Increased sliding velocity also increased the value of COF for hybrid GJFR plastic composites. It might be due to the exposure of fibres to the surface at high velocity. Exposed fibres got breakage during sliding and their worn-out detached particles led to increase COF value between the contact surfaces. The obtained results and affecting factors behind the

Figure 7 Three-dimensional response surface plot showing effect of hybrid filler Wt% and sliding velocity on COF values

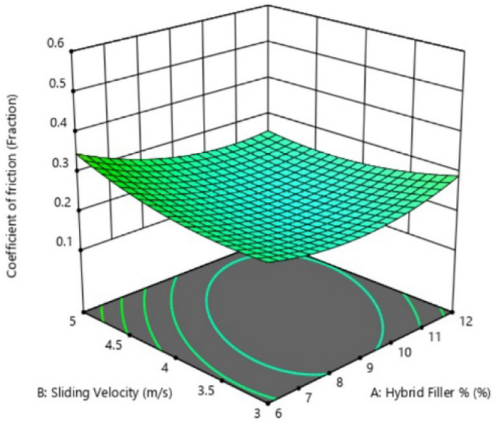
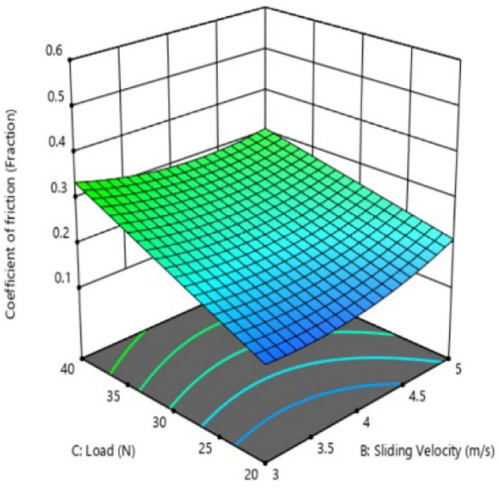


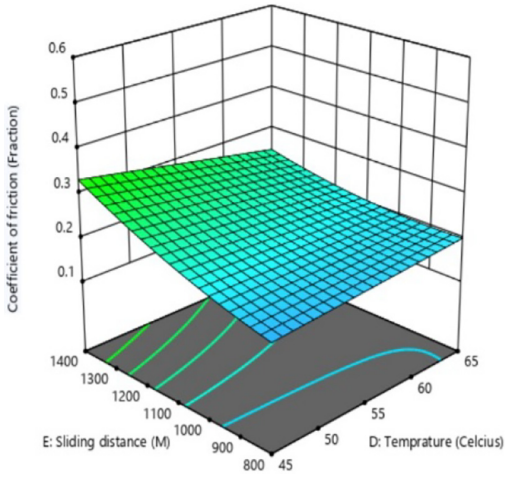
Figure 8 Three-dimensional response surface plot showing effect of load and sliding velocity on COF values



increased value of COF due to increased sliding velocity showed good agreement with the study of Ramesh and Rao (1983). The responsible factor for changed characteristics of the surface and increased COF value were found in agreement with Vishwanath *et al.* (1992). Figure 8 shows the influence of sliding velocity and load on COF value and indicated that increment in both these factors increased the value of COF. It happened due to the presence of amorphous glass and jute fibres, which invited the formation of wear debris mixed with some debonded filler particles.

The presence of wear debris on sliding track under increased velocity and high loading conditions further increased the COF value. Agrawal *et al.* (2016) have also mentioned it a general trend under such conditions. There was a mutual interaction effect of temperature and sliding distance on COF value as shown in Figure 9, which indicated that increased values of both these factors, increased COF value. This was due to the

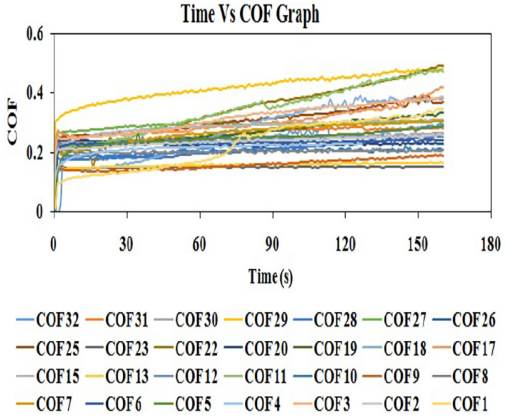
Figure 9 Three-dimensional response surface plot showing effect of temperature and sliding distance on COF values



fact that the developed composites had reinforcement of glass and jute fibres, which have more coefficient of friction as compared to the matrix of unsaturated polyester resin (Moore, 2013; Tabor, 1964; Neale, 1973). Increased temperature of counterface softened the matrix material and led to fibre exposure towards the surface. The situation increased the COF value forming wear debris for long sliding distance. Therefore, after analyzing the three-dimensional response surface plots, it can be expressed that increased values of all the variable parameters in mutual interactions like sliding velocity, load, temperature and sliding distance increased the COF value for hybrid GJFRP composites. However, the increased hybrid filler per cent was found with exceptional behaviour, which initially increased COF value and then brought a decrease in the COF value. It was happened due to the more amount of filler particles that behaved like a solid lubricant. The possible responsible factors for the variation in COF values showed agreement with the result of Suresha (2006) (Figure 10).

Time V/S COF graph indicates that initially there was a linear increment in COF values in all the experiment runs and after the formation of a thin film of matrix material on a sliding

Figure 10 Time v/s COF value graph of various experiment runs



track, the values of COF were changing slightly. Finally, there was an observed linear increment in COF values with the increment in time of sliding. There was the absence of a fixed increasing or decreasing trend in COF values with the increment in applied load values. It happened due to the relative effect of variation in the values of hybrid filler wtper cent loading, the temperature of counter-face, sliding velocity and sliding distance.

Figure 11 Hybrid GJFRP with 3 Wt% fillers loading post wear test showing wear debris, deboned fillers and worn out fibres

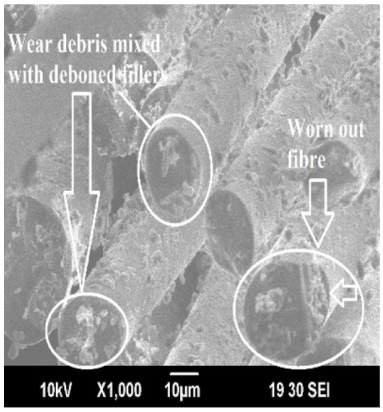


Figure 12 Hybrid GJFRP with 6 Wt% fillers loading post wear test showing worn out fibres

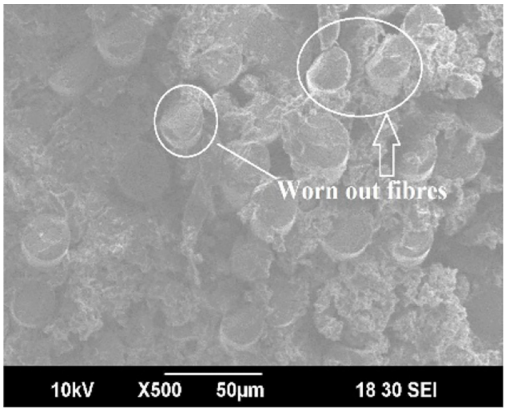


Figure 13 Hybrid GJFRP with 9 Wt% filler loading exposing damage to the matrix and wear debris

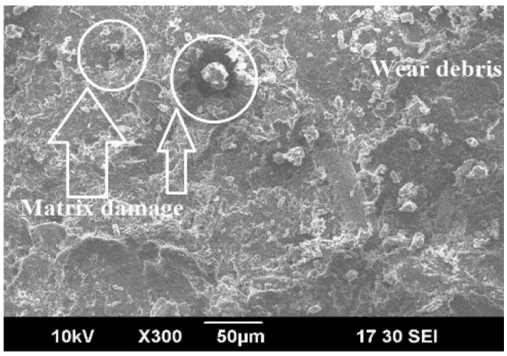


Figure 14 Hybrid GJFRP with 12 Wt% filler loading exhibiting worn out the matrix with cracks formation

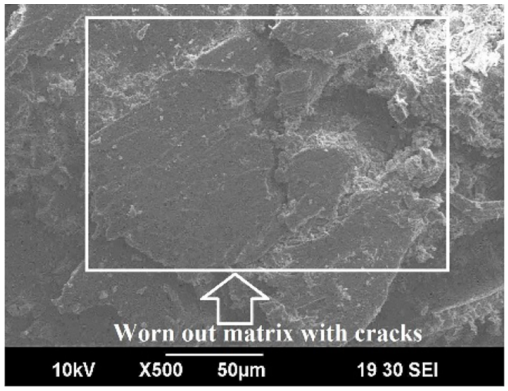
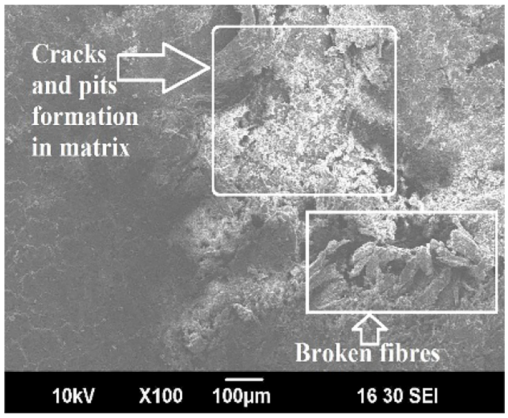


Figure 15 Hybrid GJFRP with 15 Wt% filler loading showing broken fibres and cracks and pits formation in the matrix material



SEM images in Figures 11–15, exposed the wear debris, worn out fibres, debonded fillers, damage to the matrix material, worn out matrix-material and cracks formation. Table 6 showed the values of optimized process variables to minimize the coefficient of friction and confirmation experiment results.

3.2 Wear

Table 2 shows the result of various experiments to wear values along with actual factors as per the central composite design. The analysis of variance table for wear shows the model as significant due to *F*-value of 90.83. *p*-values less than 0.05 showed significant terms of the model. The possibility of getting the larger *F*-values was 0.01% and the responsible factor for the same might be noise. From the model, terms such as A, B, C, D, E, AB, AC, AD, BD, BE, DE, A^2, B^2 and C^2 were found significant with their percentage contribution as 1.54%, 1.33%, 4.75%, 1.16%, 1.05%, 1.08%, 7.53%, 2.07%, 4.37%, 22.89%, 5.17%, 38.31%, 8.03% and 0.71%, respectively. Values having more than 0.1 gave the clue about the insignificant nature of the model terms. If a large number of

Table 6 Values of optimized process variables to minimize coefficient of friction

Optimized values of process variables									
Type of optimization	Aim	Hybrid filler %	Sliding velocity (ms^{-1})	Load (N)	Temperature (°C)	Sliding distance (m)	Response (predicted)	Response (Outcome from Experiments)	Desirability
Single response	To reduce COF	9	4	30	54	1100	0.30	0.22	0.722

such less important terms were present, then the method of model reduction can be implemented to improve the model. The results showed the *F*-value as 1.63, which represented the non-significant nature of lack of fit. In this case, there was 30.76% possibility for having a larger value of lack of fit, which might be due to noise. The predicted R^2 of 0.95 was representing a logical relationship with an adjusted R^2 of 0.98. This indicated the difference to be less than 0.2. The value of adeq precision was used to find the signal to noise ratio. Ratio value more than 4 indicates the most suitable condition and in the present model, the ratio was found 35.34 that directed towards the best suitability signal for the analysis.

The plot of actual V/S predicted values of wear indicates superb fitness of the regression model. All the outcomes were equivalent to the predicted value and for the same, this plot can be taken into consideration. The observed values were very near to the line of prediction, which indicates that all the outcomes of experimental work were in agreement with the predicted results (Figure 16).

The coded equation for wear model is as shown below:

$$\begin{aligned}\text{Wear} = {} & 11.4714 + 4.35973 * \text{A} - 4.05799 * \text{B} + 7.6556 * \text{C} \\ & - 3.79152 * \text{D} + 3.6082 * \text{E} + 4.48819 * \text{AB} \\ & + 11.8063 * \text{AC} - 6.18356 * \text{AD} - 0.990225 \\ & * \text{AE} - 1.54083 * \text{BC} + 8.99064 * \text{BD} - 20.5843 \\ & * \text{BE} - 1.39787 * \text{CD} - 0.241437 * \text{CE} - 9.78728 \\ & * \text{DE} + 19.6666 * \text{A}^2 + 9.07239 * \text{B}^2 + 2.77953 \\ & * \text{C}^2 - 0.027497 * \text{D}^2 + 1.65722 * \text{E}^2 \end{aligned} \quad (2)$$

Figure 16 Plot of actual V/S predicted values of wear

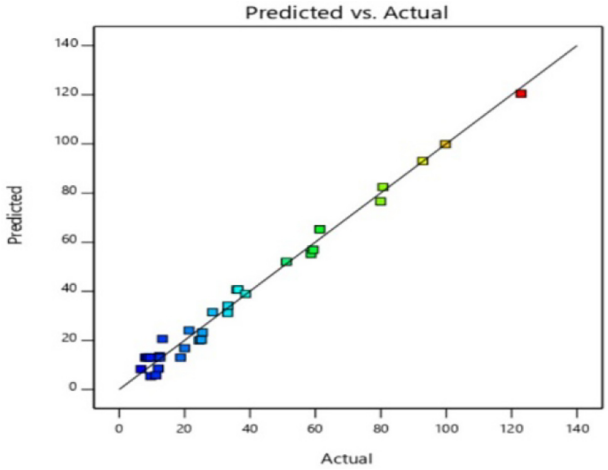


Figure 17 Three-dimensional response surface plot showing effect of hybrid filler Wt% and sliding velocity on wear values

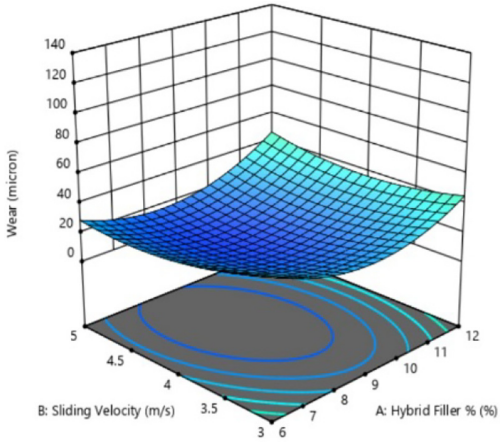
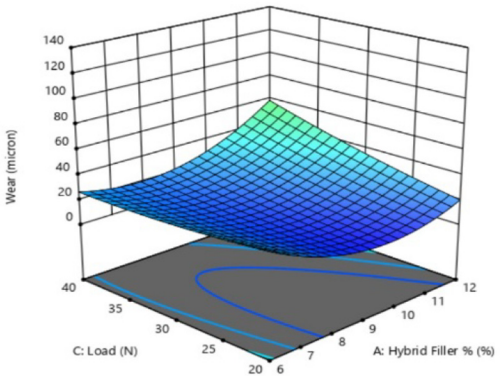


Figure 18 Three-dimensional response surface plot showing effect of hybrid filler Wt% and load on wear values



The integrated influence of various dry sliding wear test parameters on the specific wear rate of hybrid GJFRP composites has been observed from three-dimensional response surface plots (Figures 17–22). The combined effect of hybrid filler per cent and sliding velocity on wear behaviour has been explained through Figure 17 that showed similar behaviour as for COF values. Hybrid filler per cent increment firstly reduced the wear amount up to a level but after that, it increased the same. It might be due to the increased amount of filler per cent, there were also the chances of agglomeration of these particles. The clustered filler particles separated from the matrix in the form of large wear debris, which increased the wear loss (Akbarpour *et al.*, 2019). Increased sliding velocity also decreased the wear loss. Although there was an increase in temperature of contact surfaces at high sliding velocity

Figure 19 Three-dimensional response surface plot showing effect of hybrid filler Wt% and temperature on wear values

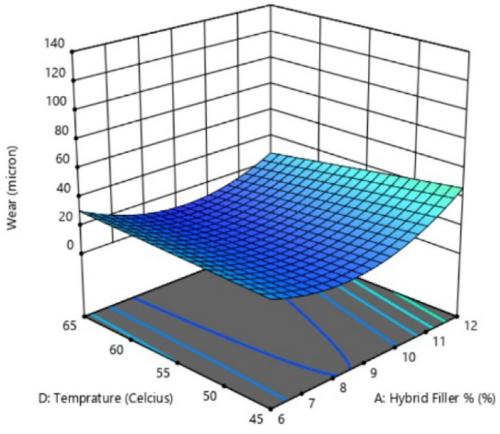


Figure 20 Three-dimensional response surface plot showing effect of sliding velocity and temperature on wear values

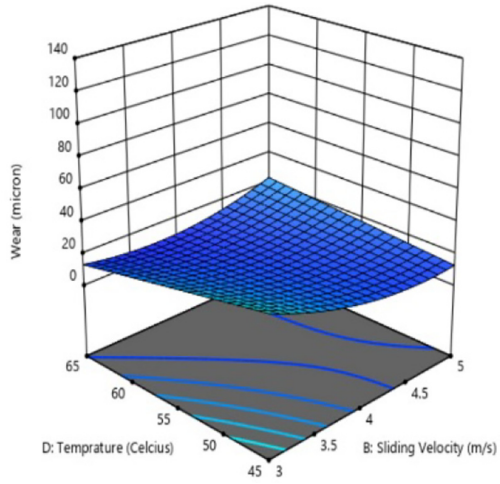


Figure 21 Three-dimensional response surface plot showing effect of sliding velocity and sliding distance on wear values

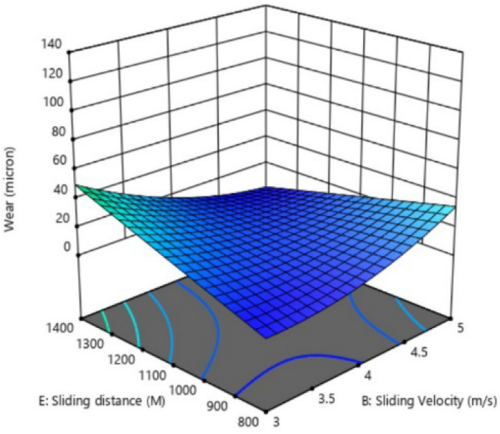
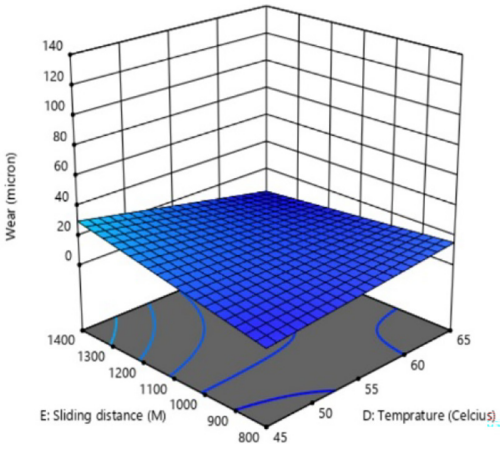


Figure 22 Three-dimensional response surface plot showing effect of sliding distance and temperature on wear values



and wear rate must have increased, but the low value of co-efficient of thermal expansion of FRP did not increase the same in spite of high sliding velocity condition. The resulted condition was in agreement with Srinivasan *et al.* (2007). Secondly, it happened due to the increment in lubricity caused by increased hybrid filler particles, which worked as solid lubricants at increased sliding velocity. Sliding velocity acted as an interacting variable parameter and found with its relative effect. Hybrid filler per cent in interaction with load decreased the wear initially but further, it raised the wear loss, which was not found correct in case of the applied load. Increment in load condition increased the wear loss. This resulted due to the fact that reduction in binding strength caused by the increase of applied load and it eventually invited more wear loss in terms of mass (Madhanagopal and Gopalakannan, 2018). Also, there was a generation of more frictional heat with the increased normal load at the contact surface and it decreased the material strength leading to more wear (Chowdhury *et al.*, 2011). There was thermal activation of crack propagation velocity at high load condition, which gave rise to a steady-state wear rate, which is in agreement with Lhymn, (1987). Effect of mutual interaction of hybrid filler per cent and temperature on wear can be observed from Figure 19, which shows that increased hybrid filler per cent decreased the wear loss initially but further observed as increasing. Temperature increment proceeded with increased wear loss. This was due to the relative effect of temperature as contact counter face temperature rise produced softening effects and hence asperities of the counter face body ploughed away the material easily from the contact surface of test specimen resulting in more amount of wear in the form of mass loss (Prabhu, 2015). Increased sliding velocity initially brought a decrease in the wear amount, after that there was a slight increase in wear. The obtained results are in agreement with the study results of Ramesh and Rao (1983). However, increased temperature reduced the wear loss as shown in Figure 20. This happened due to the presence of hybrid filler particles, which were a combination of powder form of carbon black, eggshells ash and coconut coir ash. Due to the increased temperature, these filler particles developed a very thin film between the surfaces of the pin type sample and steel disc. The chances of formation of wear rate

reducing oxide layer and glazing layer at high-temperature condition cannot be ignored. The obtained results were in agreement with the result of Rajaram *et al.* (2010). From this study, it can be concluded that melt lubrication presence at high-temperature condition reduced both friction and wear rate significantly. From Figures 21 and 22, it was observed that increase in both temperature and sliding velocity in interaction with sliding distance lead to increase in wear amount and similarly, increased sliding distance in both the cases also brought an increment in the wear. The obtained results were found in agreement with Srinivasan *et al.* (2007). This happened due to the increased sliding velocity and sliding distance, the contact time of sliding surfaces was also increased (Chowdhury and Helali, 2007). The clustered filler particles were separated from the matrix in the form of wear debris continuously for a long duration of slide thereby increasing the wear loss (Akbarpour *et al.*, 2019). There was an observed relative effect of mutual interacting parameters on wear and in order to minimize the same, process variables were optimized. Optimized values of process variables for wear minimization are shown in Table 7 along with confirmation experiment results.

Figure 23 shows wear results obtained from wear test set-up, where wear 1 value is for wear test run 1, wear 2 for wear test run 2, etc. It was observed that at higher sliding velocities ($5\,m.s^{-1} - 6\,m.s^{-1}$) and increased temperature (65°C–75°C) of counterface, wear loss was tremendous. Frictional heat generation at high sliding velocity ($5\,m.s^{-1}$ – $6\,m.s^{-1}$) deteriorated the adhesive bonding between matrix material (resin) and hybrid fibres, whereas, at moderate increased temperature range (~ 55°C), there was a considerable decrease in the wear loss as compared to a higher temperature range (65°C–75°C). Also, the surface film formation on counter-face took place. Surface film is a combined mixture of matrix material and hybrid filler particles used. Thus, the formed surface film somewhat lubricated the dry surface during the wear process and reduced the wear loss. Hence, an optimum value of the temperature of the counterface (55°C) was suggested to minimize the wear loss. Figure 24 shows wear track and disc status during wear test runs. There was a deposition of matrix material along with the minor presence of wear-out debonded filler particles forming a thin film on the disc (Agrawal *et al.*, 2016). A very high amount of hybrid filler particles loading was not considered suitable for ideal results. In case of high weight per cent loading of hybrid fillers, there was observed clustering of filler particles at certain positions in developed hybrid composite materials and it reduced the chances of filler-fibre bonding and increased the filler-filler particles bonding. As a result, these filler particles easily detached from the surface and increased the wear loss. Hence, an optimum value of hybrid filler wt per cent (9 Wt%) loading was recommended as shown in Table 7.

Figure 23 Time V/S wear loss graph for hybrid GJFRP composites

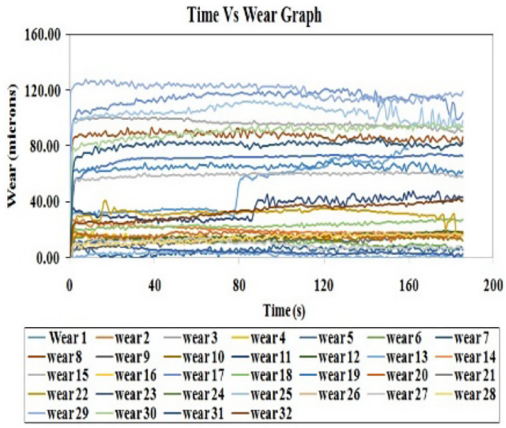
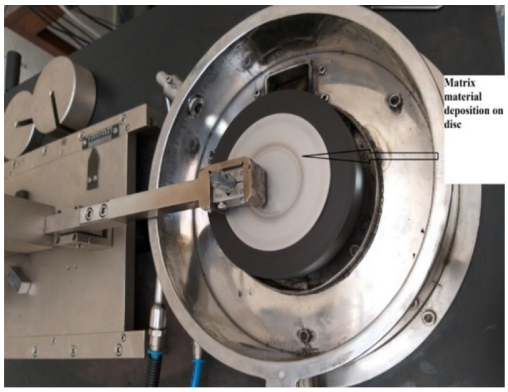


Figure 24 Wear track and disc status



The surface microstructure of the wear test samples was observed by scanning electron microscopy machine (JEOL-JSM 6390 L) after wear experiments. All the test samples were coated with a thin gold layer using a sputtering unit before the SEM study. Wear mechanism, wear tracks, debonding of fibre and matrix, cracks, pits and debris formation were observed through SEM images. Results of microscopy study were explained with the help of various SEM images as shown in Figures 25–30. The exposed array of fibres were observed in Figure 25. Exposer of fibres took place when matrix material was removed during sliding and deposited on disc forming wear track impression. Fibre breakage and wear debris were observed in Figure 26 image. Filler particles were debonded from matrix material during sliding and wear debris were formed. Pits were also formed due to loosening of matrix material and debonding of filler particles during dry sliding at an increased temperature of

Table 7 Values of optimized process variables for single response to minimize wear

		Optimized values of process variables							
Type of optimization	Aim	Hybrid filler %	Sliding velocity ($m.s^{-1}$)	Load (N)	Temperature (°C)	Sliding distance (m)	Response (predicted)	Response (Outcome from Experiments)	Desirability
Single response	To reduce wear	9	4	30	55	1,100	16.09	12.96	0.95

Figure 25 Hybrid GJFRP with 3% filler after wear

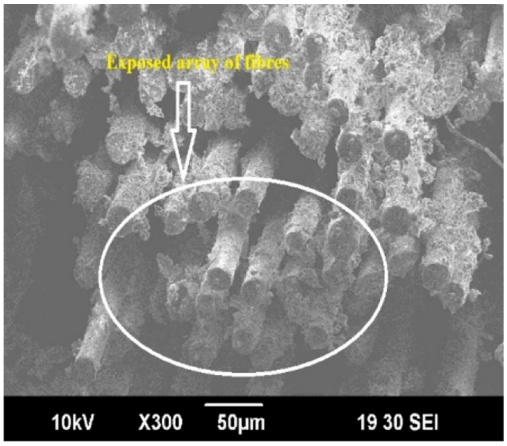


Figure 26 Hybrid GJFRP with 6% filler after wear

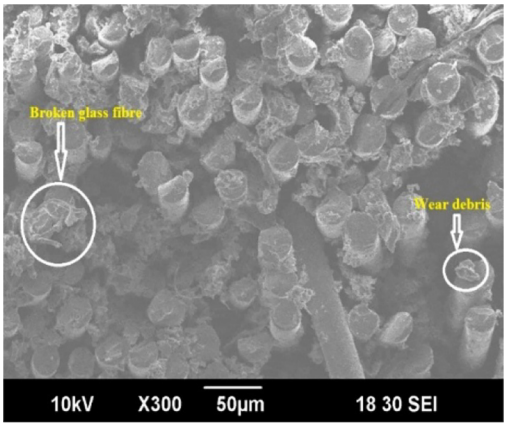


Figure 27 Hybrid GJFRP with 9% filler after wear

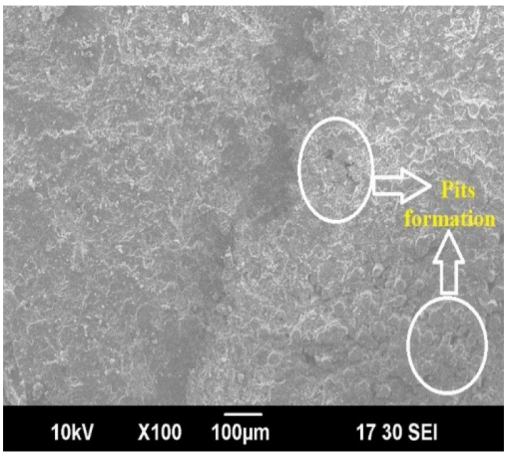


Figure 28 Hybrid GJFRP with 9% filler after wear

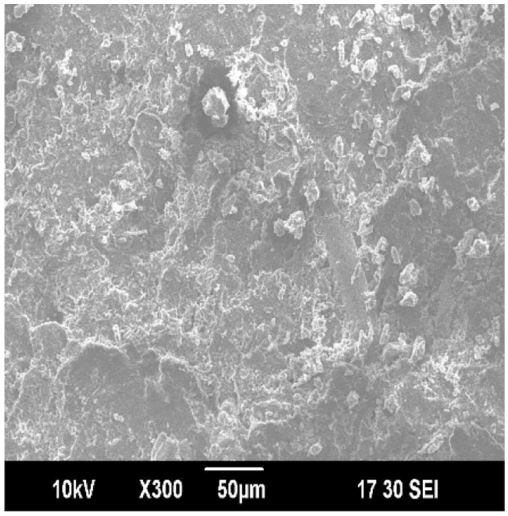


Figure 29 Hybrid GJFRP with 12% filler after wear

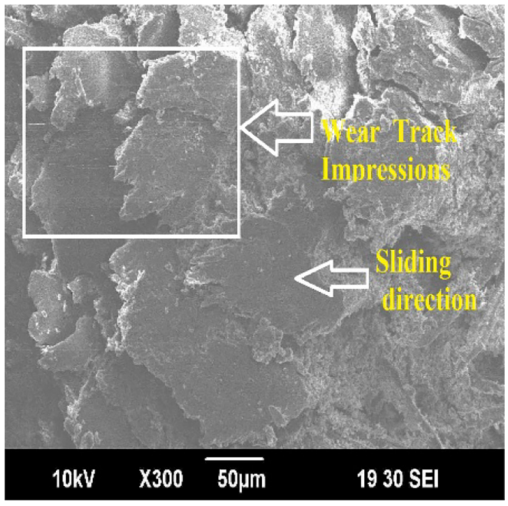
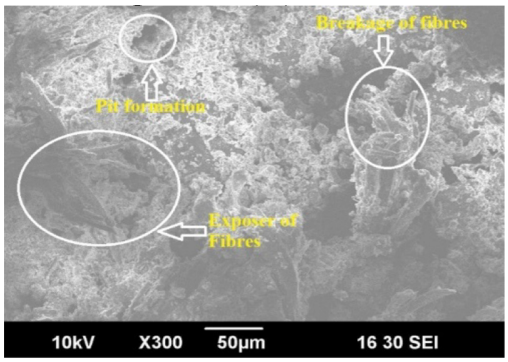


Figure 30 Hybrid GJFRP with 15% filler after wear



counterface. Wear track can be observed from Figure 29 when wear pin was sliding on the disc.

3.3 Friction force

Table 2 shows the results of various experiments for friction force values along with actual factors as per the central composite design. The analysis of variance table for friction force shows the model as significant due to *F*-value of 19.58. *p*-values less than 0.05 showed significant terms of the model. The possibility of getting the larger *F*-values is 0.01% and the responsible factor for the same might be noise. From the model terms such as A, B, C, D, E, AB, AC, BE, A^2, B^2, D^2 and E^2 were found significant with their percentage contribution as 0.23%, 15.7%, 8.69%, 1.03%, 12.86%, 18.02%, 3.7%, 12.11%, 11.54%, 10.92%, 2.12% and 3.08%, respectively. Values greater than 0.1 gave clue about the insignificant nature of the model terms. If there is a presence of a large number of such terms which are less important, then the method of model reduction can be implemented to improve the model. The results showed the *F*-value as 0.37, which represented the non-significant nature of lack of fit. In this case, there was 93.36% possibility for having a larger value of lack of fit, which might be due to noise. The predicted R^2 of 0.83 represented a logical relationship with an adjusted R^2 of 0.88. This indicated the difference of less than 0.2. The value of adeq precision was used to find the signal to noise ratio. If the ratio went beyond 4, then it is the most suitable condition. The present model has a ratio of 15.95 that indicates the best suitability signal for the analysis.

Figure 31 indicates the superb fitness of the regression model. The outcomes were equivalent to the predicted value and for the same, actual V/S predicted values of friction force plot can be taken into consideration. The observed value points were very near to the line of prediction, which indicates that the outcomes of experimental work are in agreement with the predicted results.

The coded equations for friction force model is as shown below:

$$\begin{aligned}\text{Friction force} = {} & 6.19065 + 0.210617 * A - 1.74376 * B \\ & + 1.29712 * C - 0.44595 * D + 1.57827 * E \\ & + 2.28783 * AB + 1.03704 * AC - 1.87593 \\ & * BE + 1.34788 * A^2 + 1.31075 * B^2 \\ & + 0.578102 * D^2 + 0.696452 * E^2 \end{aligned} \quad (3)$$

Figure 31 Plot of actual V/S predicted values of friction force

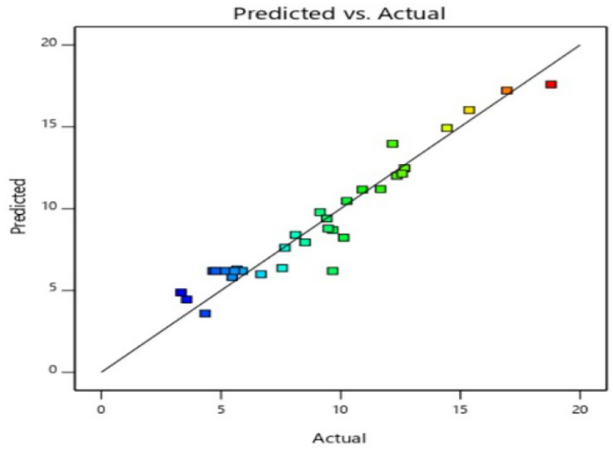


Figure 31 Actual V/S predicted values graph for friction force is as given below:

Three-dimensional response surface plots are shown in Figures 32–34, show the effect of various wear test process parameters on the magnitude of friction force for hybrid GJFRP composites. Mutual interaction of hybrid filler per cent

Figure 32 Three-dimensional response surface plot showing effect of sliding velocity and hybrid filler on friction force

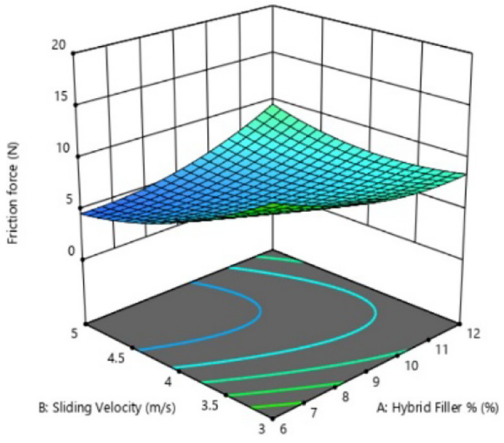


Figure 33 Three-dimensional response surface plot showing effect of hybrid filler and load on friction force

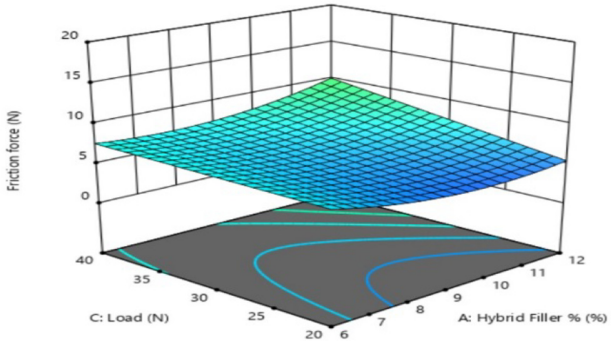
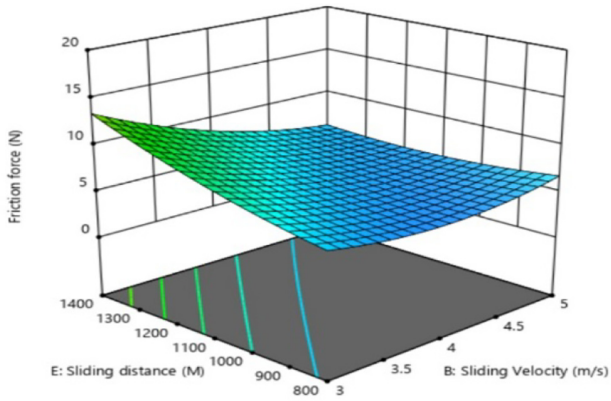


Figure 34 Three-dimensional response surface plot showing effect of sliding distance and sliding velocity on friction force



with sliding velocity as shown in Figure 32 indicated that with the increase in loading of hybrid filler per cent, friction force initially decreased and then, it increased with further increment. Increased sliding velocity had decreased the friction force amount proceeding with a slight increase. From Figure 33, it was noticed that increased hybrid filler per cent initially decreased the friction force slightly and then there was a continuous increment. Wear and friction behaviour of polymeric materials were highly dependent on their transfer behaviour on counterface material in the form of a film. The intrinsic properties of the filler particles highly influenced the transfer behaviour of such materials. These filler particles played an important role in the development of a thin, uniform and adherent transfer film to the substrate in the polymeric materials, hence they were playing a significant role in wear minimization of polymeric materials. But there was a possibility of filler particles decomposition due to the heat produced at the interface during dry sliding and consecutively formation of transition layer (Suresha *et al.*, 2010; Hsu *et al.*, 2000; Bahadur, 2000; Bahadur and Gong, 1992; Suh and Saka, 1980). Adhesion between FRP composites and the metallic disc was reduced by the formed transfer film during dry sliding and consecutively, it also diminished the dig in action by the asperities of hard steel counterface over the soft surface of polymer composites. A similar material layer interacted with fibre-matrix based generated transfer film, as it came into existence during sliding (Suh and Saka, 1980; Lancaster, 1972a, 1972b). The observed COF value was maximum (0.53 fraction) at increased sliding velocity ($5\,ms^{-1}$), and consecutively there was an increased friction force value. It happened due to the disturbance in the adhesive bonds at the interface by shear strength of FRP composite materials and it totally governed the overall friction behaviour. Generally, load increment during tribological behaviour testing of FRP materials invites an increment in friction force. In the present study also, COF values increased with the increment of applied load by keeping other variable parameters constant and as a result increment in friction force was observed. It happened because, at increased load condition, there was also an increment in the contact area of surfaces, which further increased the friction force. This was in agreement with the results obtained by Lhymn (1987). Also, it followed the AMONTONS' law cited by Bowden *et al.* (1943) in their study, which states that the frictional force is directly proportional to the applied load. Keeping all other variables constant and having the only increment in both sliding velocity and sliding distance, as a result, there is an increment in the friction force as shown in Figure 34. In the friction force case also, there were some relative effects of interacting parameters with each other and to minimize the friction force, process variables were optimized as shown in Table 8 below.

4. Conclusions

The hybrid GJFR plastic composites with hybrid fillers reinforcement were developed using the pultrusion process and were tested for tribological behaviour on pin-on-disc variable temperature based rotary tribometer. Applications of such developed hybrid composites under various wear and friction causing conditions such as dry sliding medium, different sliding velocity and distance, temperature and load conditions with optimum values were found suitable and various control factors were optimized using response surface methodology. Following conclusions are drawn from the study:

- The minimum COF value was found at 9 wt% of hybrid filler loading, 4 ms^{-1} sliding velocity, 30 N applied load, 54°C temperature and 1,100 m sliding distance values. The maximum wear resistance was observed at 9 Wt%, 4 ms^{-1}, 30 N, 55°C and 1,100 m, respectively. Required friction force was found minimum at 12 Wt %, 3.08 ms^{-1}, 20 N, 59°C and 1,100 m, respectively.
- Higher sliding velocity (4 ms^{-1}–6 ms^{-1}) and increased temperature (55°C–75°C) of counterface result in more wear as both these factors might have deteriorated the adhesive bonding between the matrix material (resin) and hybrid fibre.
- Wear volume and COF value of hybrid GJFRP were influenced by the volume, types (natural or synthetic fibre) and hybridization of fibres being loaded during hybrid FRP development due to their individual lubricating ability. Prediction of individual fibre's loading effect on tribological behaviour of FRP material is a challenging task due to their combining effects.
- Addition of glass fibres with jute fibres enhanced the wear resistance behaviour of developed hybrid composites. Hybridization of natural fibres (jute fibres) with synthetic fibres (glass fibres) improved their wear resistance strength.
- Surface morphology such as wear mechanism, wear tracks, debonding of fibre and matrix, cracks, pits and debris formation were studied through SEM images.
- Due to the light weight, low cost and good tribological behaviour, one can observe the hybrid GJFRP composites for a large number of applications such as sturdy warehouse racks in chemical industry, surface flooring, shoe soles, tent and sports use poles, moulded seat parts of buses and other body parts of automobiles, many structural applications such as walkways and bridges and many more pultruded FRP profiles.

Table 8 Values of optimized process variables for single response to minimize friction force

		Optimized values of process variables							
Type of optimization	Aim	Hybrid filler %	Sliding velocity ($m.s^{-1}$)	Load (N)	Temperature (°C)	Sliding distance (m)	Response (predicted)	Response (outcome from experiments)	Desirability
Single response	To reduce friction force	12	3.08	20	59	1,100	9.25	6.19	1.00

References

Agrawal, S., Singh, K.K. and Sarkar, P.K. (2016), "A comparative study of wear and friction characteristics of glass fibre reinforced epoxy resin, sliding under dry, oil-lubricated and inert gas environments", *Tribology International*, Vol. 96, pp. 217-224, doi: 10.1016/j.triboint.2015.12.033.

Akbarpour, M.R., Alipour, S., Safarzadeh, A. and Kim, H.S. (2019), "Wear and friction behavior of self-lubricating hybrid Cu-(SiC+ x CNT) composites", *Composites Part B: Engineering*, Vol. 158, pp. 92-101, doi: 10.1016/j.compositesb.2018.09.039.

Akpan, E.I., Wetzel, B. and Friedrich, K. (2018), "A fully biobased tribology material based on acrylic resin and short wood fibres", *Tribology International*, Vol. 120, pp. 381-390, doi: 10.1016/j.triboint.2018.01.010.

Andrich, M., Hufenbach, W., Kunze, K. and Scheibe, H.J. (2013), "Characterisation of the friction and wear behaviour of textile reinforced polymer composites in contact with diamond-like carbon layers", *Tribology International*, Vol. 62, pp. 29-36, doi: 10.1016/j.triboint.2013.01.023.

Bahadur, S. (2000), "The development of transfer layers and their role in polymer tribology", *Wear*, Vol. 245 Nos 1/2, pp. 92-99, doi: 10.1016/S0043-1648(00)00469-5.

Bahadur, S. and Gong, D. (1992), "The role of copper compounds as fillers in the transfer and wear behavior of polyetheretherketone", *Wear*, Vol. 154 No. 1, pp. 151-165, doi: 10.1016/0043-1648(92)90251-3.

Bowden, F.P., Moore, A.J.W. and Tabor, D. (1943), "The ploughing and adhesion of sliding metals", *Journal of Applied Physics*, Vol. 14 No. 2, pp. 80-91, doi: 10.1063/1.1714954.

Chairman, C.A. and Kumaresh Babu, S.P. (2013), "Mechanical and abrasive wear behavior of glass and basalt fabric-reinforced epoxy composites", *Journal of Applied Polymer Science*, Vol. 130 No. 1, pp. 120-130, doi: 10.1002/app.39154.

Chauhan, S.R., Kumar, A., Singh, I. and Kumar, P. (2010), "Effect of fly ash content on friction and dry sliding wear behavior of glass fiber reinforced polymer composites – a Taguchi approach", *Journal of Minerals and Materials Characterization and Engineering*, Vol. 9 No. 4, pp. 365-387, doi: 10.4236/jmmce.2010.94027.

Chowdhury, M.A. and Helali, M.M. (2007), "The effect of frequency of vibration and humidity on the wear rate", *Wear*, Vol. 262 Nos 1/2, pp. 198-203, doi: 10.1016/j.wear.2006.05.007.

Chowdhury, M.A., Khalil, M.K., Nuruzzaman, D.M. and Rahaman, M.L. (2011), "The effect of sliding speed and normal load on friction and wear property of aluminum", *International Journal of Mechanical & Mechatronics Engineering*, Vol. 11 No. 1, pp. 45-49, doi: 111701-6868, IJMME-IJENS © February 2011 IJENS

Dalbehera, S. and Acharya, S.K. (2015), "Effect of cenosphere addition on erosive wear behaviour of jute-glass reinforced composite using taguchi experimental design", *Materials Today: Proceedings*, Vol. 2 Nos 4/5, pp. 2389-2398, doi: 10.1016/j.matpr.2015.07.176.

Giltrow, J.P. and Lancaster, J.K. (1970), "The role of the counterface in the friction and wear of carbon fibre reinforced thermosetting resins", *Wear*, Vol. 16 No. 5, pp. 359-374, doi: 10.1016/0043-1648(70)90102-X.

Gupta, A., Singh, H. and Walia, R.S. (2015), "Effect of fillers on tensile strength of pultruded glass fiber reinforced polymer composite", availabel at: http://nopr.niscair.res.in/handle/123456789/31247

Gupta, A., Singh, H. and Walia, R.S. (2016), "Hybrid filler composition optimization for tensile strength of jute fibre-reinforced polymer composite", *Bulletin of Materials Science*, Vol. 39 No. 5, pp. 1223-1231, doi: 10.1007/s12034-016-1248-1.

Gupta, A., Singh, H. and Walia, R.S. (2014), "Effect of glass fiber and filler volume fraction variation on mechanical properties of GFRP composite", *Proceedings of the International Conference on Research and Innovations in Mechanical Engineering* (pp. 407-414). Springer, New Delhi, doi: 10.1007/978-81-322-1859-3_38

Gupta, A., Vaishya, R., Khan, K.L.A., Walia, R.S. and Singh, H. (2019), "Multi-response optimization of hybrid filler composition for pultruded jute fiber reinforced polymer composite", *Materials Research Express*, Vol. 6 No. 11, p. 115324, doi: 10.1088/2053-1591/ab4945.

Gupta, A., Vaishya, R., Khan, K.L.A., Walia, R.S. and Singh, H. (2019), "Multi-response optimization of the mechanical properties of pultruded glass fiber composite using optimized hybrid filler composition by the gray relation grade analysis", *Materials Research Express*, Vol. 6 No. 12, p. 125322, doi: 10.1088/2053-1591/ab56f2.

Hsu, D.K., Barnard, D.J., Peters, J.J. and Dayal, V. (2000), "Physical basis of tap test as a quantitative imaging tool for composite structures on aircraft", *AIP Conference Proceedings*, May, American Institute of Physics, Vol. 509, No. 1, pp. 1857-1864, doi: 10.1063/1.1306256.

Joshi, S.V., Drzal, L.T., Mohanty, A.K. and Arora, S. (2004), "Are natural fiber composites environmentally superior to glass fiber reinforced composites?", *Composites Part A: Applied Science and Manufacturing*, Vol. 35 No. 3, pp. 371-376, doi: 10.1016/j.compositesa.2003.09.016.

Karthikeyan, S., Rajini, N., Jawaid, M., Winowlin Jappes, J.T., Thariq, M.T.H., Siengchin, S. and Sukumaran, J. (2017), "A review on tribological properties of natural fiber based sustainable hybrid composite", *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, Vol. 231 No. 12, pp. 1616-1634, doi: 10.1177/1350650117705261.

Kaya, İ. and Parlar, Z. (2020), "The investigation of tribological behavior of carbon fiber-reinforced composite materials", *Industrial Lubrication and Tribology*, Vol. 72 No. 2, pp. 211-216, doi: 10.1108/ILT-05-2018-0184.

Koronis, G., Silva, A. and Fontul, M. (2013), "Green composites: a review of adequate materials for automotive applications", *Composites Part B: Engineering*, Vol. 44 No. 1, pp. 120-127, doi: 10.1016/j.compositesb.2012.07.004.

Kumar, N., Walia, R.S. and Angra, S. (2021), "Study of mechanical properties of pultruded jute-glass reinforced unsaturated polyester bio-composites with hybrid filler loading", *World Journal of Engineering*, doi: 10.1108/WJE-04-2020-0127.

Lancaster, J.K. (1972), "Friction and wear", *Polymer Science*, pp. 959-1046, Elsevier, doi: 10.1016/B978-0-7204-0247-6.50005-6

Lancaster, J.K. (1972), "Polymer-based bearing materials: the role of fillers and fibre reinforcement", *Tribology*, Vol. 5 No. 6, pp. 249-255, doi: 10.1016/0041-2678(72)90103-0.

Latha, P.S., Rao, M.V., Kumar, V.K., Raghavendra, G., Ojha, S. and Inala, R. (2016), "Evaluation of mechanical and tribological properties of bamboo–glass hybrid fiber

reinforced polymer composite", *Journal of Industrial Textiles*, Vol. 46 No. 1, pp. 3-18, doi: 10.1177/1528083715569376.
Lhymn, C. (1987), "Effect of normal load on the specific wear rate of fibrous composites", *Wear*, Vol. 120 No. 1, pp. 1-27, doi: 10.1016/0043-1648(87)90130-X.
Mangalaraja, R.V., Raj, T.M. and Prabha, D.R. (2019), Statistical features of epoxy resin based hybrid composites reinforced with jute, banana and flax natural fibers, *Journal of Chemical Technology & Metallurgy*, Vol. 54 No. 1.
Madhanagopal, A. and Gopalakannan, S. (2018), "Dry slide wear behavior of boron carbide filled epoxy composites", *Materials Today: Proceedings*, Vol. 5 No. 2, pp. 7289-7295, doi: 10.1016/j.matpr.2017.11.397.
Monfared, R.M., Ayatollahi, M.R. and Isfahani, R.B. (2018), "Synergistic effects of hybrid MWCNT/nanosilica on the tensile and tribological properties of woven carbon fabric epoxy composites", *Theoretical and Applied Fracture Mechanics*, Vol. 96, pp. 272-284, doi: 10.1016/j.tafmec.2018.05.007.
Moore, D.F. (2013), Principles and Applications of Tribology: Pergamon International Library of Science, Technology, Engineering and Social Studies: International Series in Materials Science and Technology, Elsevier, Vol. 14, available at: www.abebooks.com/servlet/BookDetailsPL?bi=30132695341&searchurl=sortby%3D17%26tn%3Dprinciples%2Btribology&cm_sp=snippet-_-srp1-_-title4
Nanda, B.P. and Satapathy, A. (2020), "An analysis of the sliding wear characteristics of epoxy-based hybrid composites using response surface method and neural computation", *Journal of Natural Fibers*, pp. 1-15, doi: 10.1080/15440478.2020.1722781.
Neale, M.J. (1973), 1973. "Tribology handbook Butterworths", available at: www.abebooks.co.uk/servlet/BookDetailsPL?bi=30335785461&searchurl=isbn%3D9780470630815%26sortby%3D17&cm_sp=snippet-_-srp1-_-title1
Pati, P.R. and Satpathy, M.P. (2021), "Effect of process parameters on sliding wear performance of red brick dust-filled glass–epoxy composites", *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, p.13506501211010553, doi: 10.1177/13506501211010553
Prabhu, T.R. (2015), "Effects of solid lubricants, load, and sliding speed on the tribological behavior of silica reinforced composites using design of experiments", *Materials & Design*, Vol. 77, pp. 149-160, doi: 10.1016/j.matdes.2015.03.059.
Rajaram, G., Kumaran, S., Rao, T.S. and Kamaraj, M. (2010), "Studies on high temperature wear and its mechanism of Al–Si/graphite composite under dry sliding conditions", *Tribology International*, Vol. 43 No. 11, pp. 2152-2158, doi: 10.1016/j.triboint.2010.06.004.
Ramesh, R. and Rao, R.M.V.G.K. (1983), "Dry wear studies on glass-fibre-reinforced epoxy composites", *Wear*, Vol. 89 No. 2, pp. 131-136, doi: 10.1016/0043-1648(83)90238-7.
Ruggiero, A., Merola, M., Carlone, P. and Archodoulaki, V.M. (2015), "Tribo-mechanical characterization of reinforced epoxy resin under dry and lubricated contact conditions", *Composites Part B: Engineering*, Vol. 79, pp. 595-603, doi: 10.1016/j.compositesb.2015.05.015.
Srinivasan, V., Maheshkumar, K.V., Karthikeyan, R. and Palanikumar, K. (2007), "Application of probablistic neural network for the development of wear mechanism map for glass fiber reinforced plastics", *Journal of Reinforced Plastics and Composites*, Vol. 26 No. 18, pp. 1893-1906, doi: 10.1177/0731684407082632.
Suh, N.P. and Saka, N. (1980), Fundamentals of tribology (No. CONF-780666-), The MIT Press, Cambridge, MA, available at: https://mitpress.mit.edu/books/fundamentals-tribology
Suresha, B. (2006), "The role of fillers on friction and slide wear characteristics in glass-epoxy composite systems", *Journal of Minerals and Materials Characterization and Engineering*, Vol. 5 No. 1, p. 87, doi: 10.4236/jmmce.2006.51006.
Suresha, B., Kumar, K.S., Seetharamu, S. and Kumaran, P.S. (2010), "Friction and dry sliding wear behavior of carbon and glass fabric reinforced vinyl ester composites", *Tribology International*, Vol. 43 No. 3, pp. 602-609, doi: 10.1016/j.triboint.2009.09.009.
Tabor, D. (1964), *The Friction and Lubrication of Solids*, Clarendon Press.
Tripathi, V.K. and Ambekar, S. (2019), "Optimization and analysis of wear rate of CFRP-NanoZno/nanoclay hybrid composites using RSM", *Journal of Bio-and Tribo-Corrosion*, Vol. 5 No. 4, pp. 1-13, doi: 10.1007/s40735-019-0284-y.
Vishwanath, B., Verma, A.P. and Rao, C.K. (1992), "Effect of matrix content on strength and wear of woven roving glass polymeric composites", *Composites Science and Technology*, Vol. 44 No. 1, pp. 77-86, doi: 10.1016/0266-3538(92)90027-Z.
Yallew, T.B., Kumar, P. and Singh, I. (2014), "Sliding wear properties of jute fabric reinforced polypropylene composites", *Procedia Engineering*, Vol. 97, pp. 402-411, doi: 10.1016/j.proeng.2014.12.264.

Further reading

Kumar, N., Angra, S. and Walia, R.S. (2019), "Development and comparison of tensile and compressive strength and percentage shrinkage of glass-jute hybrid fibre reinforced polymer composites", *Journal of Physics: Conference Series*, Vol. 1240 No. 1, p. 012123, doi: 10.1088/1742-6596/1240/1/012123, July, IOP Publishing.
Lim, S.C. and Ashby, M.F. (1987), "Overview no. 55 wear-mechanism maps", *Acta Metallurgica*, Vol. 35 No. 1, pp. 1-24, doi: 10.1016/0001-6160(87)90209-4.
Standard, A.S.T.M., 2010. G99-05 (2010), *Standard Test Method for Wear Testing with a Pin-on-Disk Apparatus*, ASTM International, West Conshohocken, PA.

About the authors



Navin Kumar: Mr Navin Kumar, born in 1986, is currently working as an Assistant Professor in the mechanical engineering department, Greater Noida Institute of Engineering and Technology, Greater Noida, India and research scholar at Mechanical Engineering Department, National Institute of Technology, Kurukshetra, India. He has more than ten years of teaching experience. He completed his BTech from Kurukshetra

University, Kurukshetra, India and MTech from M.D. University, Rohtak, India. His area of research is advanced materials and its characterization, advanced machining methods etc. Navin Kumar is the corresponding author and can be contacted at: navin4240@gmail.com



Prof R.S. Walia: Professor (Dr) Ravinderjeet Singh Walia is currently working as a Professor and head in Production and Industrial engineering department, Punjab Engineering College (Deemed to be University), Chandigarh, India. Earlier he was a Professor in Delhi Technological University, Delhi. He completed his PhD from IIT, Roorkee. He has more than 115 publications in various international journals of repute and completed 5 technical projects successfully. He has supervised fifteen PhD students till now. His area of research is advanced manufacturing processes, Production and industrial engineering, Human factor engineering and FEM.



Prof Surjit Angra: Professor (Dr) Surjit Angra, born in 1960, is currently working as Senior Professor in Mechanical Engineering Department, National Institute of Technology, Kurukshetra, India. He completed his PhD in 1995 from REC, Kurukshetra (presently NIT, Kurukshetra). He has supervised 14 PhD students and supervised 45 MTech dissertations. He has published more than 100 research papers in international journals such as *Tribology Letters*, *Tribology International and Journal of ASME*. His area of research is Tribology and Machine Design.

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**

2021 Fourth International Conference on Computational Intelligence and Communication Technologies (CCICT) | 978-1-6654-2392-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/CCICT53244.2021.00028

Turing Machine based Syllable Splitter

Abhishek Mallik
Dept. of Computer Science and Engineering
Delhi Technological University
New Delhi, India
abhishekmallik265@gmail.com

Anavi Khetarpal
Dept. of Computer Science and Engineering
Delhi Technological University
New Delhi, India
anavikhetarpal@gmail.com

***Abstract*—Nowadays, children, teens, and almost everyone around us tend to receive abundant and frequent advice regarding the usefulness of syllabification. Not only does it improve pronunciation, but it also makes it easier for us to read unfamiliar words in chunks of syllables rather than reading them all at once. Within this paper, we have designed, implemented, and presented a Turing machine-based syllable splitter. A Turing machine forms the theoretical basis for all modern computers and can be used to solve universal problems. On the other hand, a syllable splitter is used to hyphenate words into their corresponding syllables. We have proposed our work by illustrating the various states of the Turing machine, along with the rules it abides by, its machine specifications, and transition function. In addition to this, we have implemented a Graphical User Interface to stimulate our Turing machine to analyze our results better.**

***Keywords*-Turing Machine, Syllable Splitter, Finite Automata, Computational Theory, Modelling and Simulation**

A Turing machine is a mathematical model equivalent to a digital computer and is currently considered a foundational model of computer science and computability. It was invented by the English mathematician Alan Turing in 1936 and has been widely used since then. According to Turing's hypothesis,a function is algorithmically computable only if computable by a Turing machine. A problem that a Turing machine cannot solve is incapable of being solved.

The machine accepts languages generated by type 0 grammars and comprises a one-dimensional, infinite length tape as temporary storage. The tape is divided into cells containing either a special symbol called blank or an input symbol. Associated with the tape is a read-write head pointer which can traverse in both directions and point to the cell currently being read. A Turing machine is an automatic machine; therefore, at any stage of computation, its instantaneous description can be determined by the symbol being scanned, the machine's internal state, and the reading head's location. Furthermore, it can mathematically be represented as a 7-tuple where:

Q, is a finite set of states
Γ, is the tape alphabet
Σ, is the input alphabet
δ, is a transition function
q0, is the initial state
#, is the blank symbol
F, is the set of final states

Through this paper, we aim to implement a syllable splitter by using a Turing machine. A syllable splitter is used to split words into syllables by following a set of rules. Syllables are often considered to be the phonological building blocks of words. Thus, the syllable splitter plays a crucial role in reading, spelling, and pronouncing words correctly. Syllabification divides the word into chunks of syllables, making it simple for the human brain to process it. This is because we holdonto these chunks in our short-term memory while processing and then recognizing the word. It is also easier to spell a long word more accurately if you spell it in chunks of syllables rather than attempting to recall all of its letters separately. In addition to this, when the word cannot fit on the remainder of the line, some applications find it difficult to justify the text. This results in a phenomenon called "overfull box." Here, the syllable splitter can be used to hyphenate words at the end of the line.

The rest of the paper is organised as follows: Section I presents the previously done work in the underlying field. Section II illustrates our proposed work in great depth. Experimental results and its analysis is present in section III. While the concluding remarks and the scope for future work is given in section IV.

I. Related Work

After the invention of the Turing machine in 1936 [1], it has been increasingly used to implement mathematical operations along with a variety of other problems. Researchers in this field have been concentrating on designs as well as attempting to lay down its practical applications. Researchers Dutra et al. [2] in their works, have designed and implemented a Blind Turing Machine to perform secure function evaluation (SFE) by using group homomorphic encryption. They have also evaluated the Turing machine implementation with differing degrees of security guarantee and compared it to a garbled- circuit-based implementation. Saqib et al. [3] performed a study on encryption and decryption using automata theory. Their primary aim was securing data in the communication between two parties.

978-1-6654-2392-2/21/$31.00 ©2021 IEEE
DOI 10.1109/CCICT53244.2021.00028

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:06:13 UTC from IEEE Xplore. Restrictions apply.

They have discussed the already existing work in this field and have also presented numerous encryption and decryption schemes with the help of a Turing machine.

Researchers Goldwasser et al. [4] construct cryptographic schemes for running Turing machines on encrypted data in their works. Their construction removes the previous restrictions on the homomorphic encryption scheme for Turing machines. The theorems presented make the use of Input-specific running times and avoid the worst-case running time. Researchers Tripathi et al. [5], in their study, proposed a notion that makes the use of a Turing machine to calculate the next successor decimal number from a given decimal number. They also presented algorithms that use this number generation approach in cryptographic techniques where the number generated is applied for data security enhancement. Tkacik et al. [6] drew a comparison between the properties of Neural Turing Machine and Stack RNN. They applied their findings for the sequential learning of human mobility patterns and observed that the Neural Turing Machine based predictor outperformed other predictors.

II. Proposed Work

In this work, we propose a Turing machine-based simulation of a syllable splitter. It works by using a single-tape Turing machine. It reads any given English word and processes it by iterating over every alphabet starting from the first alphabet. While iterating over the word, it scans every alphabet and analyses the nature of the alphabet and a fixed length of previously occurred alphabets. The syllable splitter then decides by considering these factors as to what decision is to be made. The decision is a combination of the choice of the next state, the direction of pointer movement, and whether or not the current alphabet is modified. This way, the proposed syllable splitter goes on splitting the word into its corresponding syllables by adding a '-' in between the syllables. It also accepts or rejects a word based on whether the given word can be appropriately split into its corresponding syllables while adhering to the given rules. It can be used to split English words into their corresponding syllables and is also useful in understanding how the hyphenation of words at the end of sentences is implemented. The rules followed while splitting the word, the machine specifications, various machine states, and the transition function for the Turing machine is specified as follows:

A. Rules of splitting syllables

For designing any Turing machine, we need a set of stipulated rules that determine whether a string should be accepted or rejected and, if it is accepted, how should it be processed. For determining the rules for our ad- hoc Turing machine, we take inspiration from the way Turkish words are split into syllables [7]. These rules can be listed as follows:

(i) A consonant between two vowels forms a syllable with the next vowel.
(ii) For two consecutive consonants between two vowels, first consonant forms a syllable with the previous vowel while the second consonant forms a syllable with the next vowel.
(iii) For three consecutive vowels between two vowels, the last consonant forms a syllable with the latter vowel while the first two consonants form a syllable with the former vowel.

Based on these rules for splitting the Turkish words, we formulate our own set of rules for English words. In addition to this, we also propose that words with two consecutive vowels can't be split into syllables. The rules that are used to split an English word into its corresponding syllables can be listed as:

(i) For a consonant between two vowels, consonant has a syllable with next vowel. vcv : v-cv.
(ii) For two consonants between vowels, one has syllable with previous vowel while the other has syllable with latter vowel. vccv : vc-cv.
(iii) For three consonants between vowels, last one has syllable with next vowel. vcccv : vcc-cv.
(iv) Anything else is rejected. cvvc : Rejected.

Here,
c is a consonant
v is a vowel.

B. Machine Specifications

Based on the above rules, a Turing machine can be defined as follows :

Q: {q0, q1, q2, q3, q4, q5, q6, q7, q8}.
Σ : {Consonants} U {Vowels}.
Γ : {Consonants } U {Vowels} U {'#'} U {'-'}.
δ : Transition Function.
q0: is the initial state.
#: is the blank symbol.
F: {q7 (Accept), q8 (Reject), Final States}.

C. Various states

The proposed Turing machine has nine states. These states are described as :

- q0 : Initial state
- q1 : ends on a vowel
- q2 : ends on a consonant
- q3 : ends on a vowel-consonant
- q4 : ends on a vowel-consonant-consonant
- q5 : ends on a vowel-consonant-consonant-consonant
- q6 : a syllable substructure is found
- q7 : Accept state
- q8 : Reject state

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:06:13 UTC from IEEE Xplore. Restrictions apply.

D. Transition Function

The transition function for the proposed Turing machine is as follows :

Table I
TRANSITION TABLE

Q\	Consonants	Vowels	#
q0	q2,c,R	q1,v,R	q8
q1	q3,c,R	q8	q7
q2	q2,c,R	q1,v,R	q7
q3	q4,c,R	q6,v,L	q7
q4	q5,c,R	q6,v,L	q7
q5	q2,c,R	q6,v,L	q7
q6	q6,-c,R	q1,v,R	q7
q7 (Accept)	-	-	-
q8 (Reject)	-	-	-

III. RESULTS

The Turing machine-based syllable splitter can be simulated using a Graphical User Interface (GUI) to assess better and analyze the words' processing. In our work, we have also implemented a GUI to visualize the Turing machine-based syllable splitter.

A. Syllable Splitter

Figure 1, shows our proposed derivation of a generic Turing machine. The purple arrows show the transitions when a consonant is encountered. The red arrows point to the transitions related to the occurrence of a vowel. While the green arrow shows the occurrence of a '#' symbol.

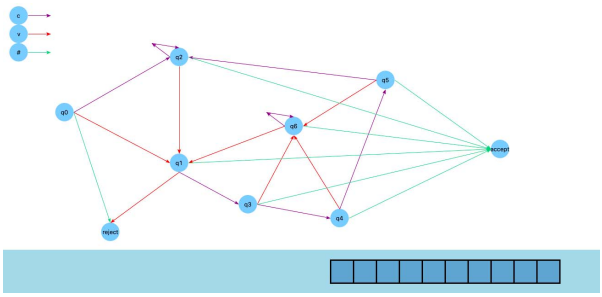


Figure 1. Implemented Turing Machine.

B. Accepted word & splitted into syllables

Figure 2, shows a word "Democracy" being accepted and splitted into its corresponding syllables. Since this word can be splitted, it is accepted by our machine. The various transition states transitioned by the turing machine are as follows:

$q0 => q2 => q1 => q3 => q6 => q6 => q1 => q3 => q4 => q6 => q6 => q1 => q3 => q4 => Accept(q7)$

While the tape symbols are read and written as follows:

$D => E => M => O => -M => O => C => R => A => -R => A => C => Y$

Here it can be seen that all the tape symbols are read and some of them are modified to highlight the splitting of the word. After processing all the tape symbols the machine enters an accept state and halts.

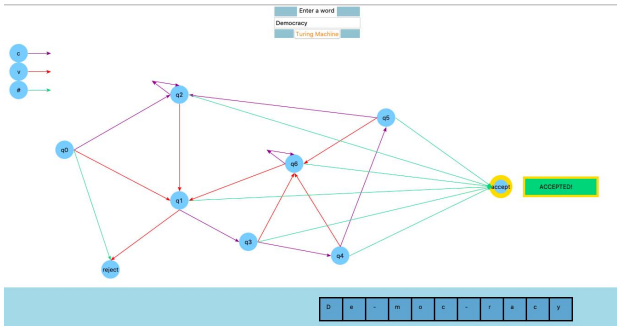


Figure 2. Accepted Word.

C. Rejected word

Figure 3, shows a word "Bear" being rejected according to the aforementioned rules. Since this word can't be accepted based on the rules for splitting the word so it is rejected by our machine. The various transition states transitioned by the turing machine are as follows:

$q0 => q2 => q1 => Reject(q8)$

While the tape symbols are read and written as follows:

$B => E => A$

Here it can be seen that all the tape symbols are not read. This happens as the string is rejected after encountering two consecutive vowels ('EA'). This causes the machine to go into reject state and halt.

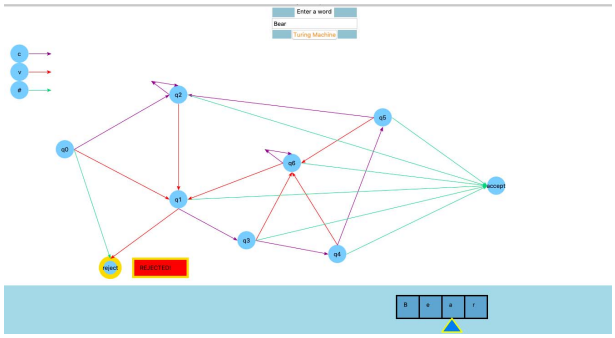


Figure 3. Rejected Word.

IV. CONCLUSION

The Turing machine is a highly powerful model of computation, and it can be used to solve any universal problem

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:06:13 UTC from IEEE Xplore. Restrictions apply.

mathematically. We have implemented a syllable splitter by using a Turing machine and abiding by a set of hyphenation rules. Our work is based on the splitting of words into syllables to make pronunciation clearer as well as increase the ability to spell and read words correctly. The splitter can also be used to divide words at the end of a line.

Even though the Turing machine is wholesome in computing problems and its computing capability can't be improved any further, we can still improve its computing speed by adding memory stacks, multiple-input tapes, etc. As part of the future work, we would like to use the Turing machine for testing the validity of grammar and for languages that can be expressed in Chomsky Normal Form (CNF). English mnemonics such as "I before e, except after c," binary computations and some other trivial mathematical problems can also be solved by implementing a suitable Turing machine. We can also attempt to show that for every Turing machine T, we can construct a context-free language (CFL) whose complement is a complete description of all the terminating computations performed by T. Other applications of the Turing machine, such as function computer and enumerator, can also be implemented.

REFERENCES

[1] Turing, A.M., 1936–7, "On Computable Numbers, With an Application to the Entscheidungsproblem", Proceedings of the London Mathematical Society, s2-42: 230–265; correction ibid., s2-43: 544–546 (1937). doi:10.1112/plms/s2-42.1.230 and doi:10.1112/plms/ s2-43.6.544

[2] Dutra, R. (2014). BlindTM – A Turing Machine System for Secure Function Evaluation.

[3] Sharipbay, Altynbek & Saukhanova, Zh & Shakhmetova, Gulmira & Saukhanov, N.. (2019). Application of finite automata in cryptography. ICEMIS '19: Proceedings of the 5th International Conference on Engineering and MIS. 1-3. 10.1145/3330431.3330452.

[4] Goldwasser S., Kalai Y.T., Popa R.A., Vaikuntanathan V., Zeldovich N. (2013) How to Run Turing Machines on Encrypted Data. In: Canetti R., Garay J.A. (eds) Advances in Cryptology – CRYPTO 2013. CRYPTO 2013. Lecture Notes in Computer Science, vol 8043. Springer, Berlin, Heidelberg.

[5] H. Tripathi and B. Hazela, "Data Security Enhancement Through Number System: A Turing-Machine Approach," 2016 Second International Conference on Computational Intelligence & Communication Technology (CICT), Ghaziabad, 2016, pp. 632-637, doi: 10.1109/CICT.2016.131.

[6] J. Tkacík and P. Kordík, "Neural Turing Machine for sequential learning of human mobility patterns," 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, 2016, pp. 2790-2797, doi: 10.1109/IJCNN.2016.7727551.

[7] http://www.tdk.gov.tr/index.php?id=208:Hece..option=com_content

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 31,2021 at 07:06:13 UTC from IEEE Xplore. Restrictions apply.

2021 ASEE ANNUAL CONFERENCE
Virtual Meeting | July 26–29, 2021 | Pacific Daylight Time
Paper ID #33632

Using the Boeing Max Air Disaster as A Role-play Scenario for Teaching Ethical Thinking

Ashish Hingle, George Mason University

Ashish Hingle (he/his/him) is a Ph.D. student in the College of Engineering & Computing at George Mason University. His research interests include engineering ethics, information systems, and student efficacy challenges in higher education. Ashish graduated with a Bachelor of Science in Computer Information Systems and a Master of Science in Information Assurance and Security from Cal Poly Pomona, California.

Dr. Aditya Johri, George Mason University

Aditya Johri is Professor in the department of Information Sciences & Technology. Dr. Johri studies the use of information and communication technologies (ICT) for learning and knowledge sharing, with a focus on cognition in informal environments. He also examine the role of ICT in supporting distributed work among globally dispersed workers and in furthering social development in emerging economies. He received the U.S. National Science Foundation's Early Career Award in 2009. He is co-editor of the Cambridge Handbook of Engineering Education Research (CHEER) published by Cambridge University Press, New York, NY. Dr. Johri earned his Ph.D. in Learning Sciences and Technology Design at Stanford University and a B.Eng. in Mechanical Engineering at Delhi College of Engineering.

Huzefa Rangwala
Dr. Alexander Monea, George Mason University

Alexander Monea is an Assistant Professor in the English Department and Cultural Studies Program at George Mason University. His research focuses on data ethics and the cultural impact of computation. He also serves as an Area Editor on Ethics, Equity, and Trust at the journal Data & Policy.

©American Society for Engineering Education, 2021

Using the Boeing Max Air Disaster as A Role-Play Scenario for Teaching Ethical Thinking

Abstract
Algorithms are a central component of most services we use across a range of domains. These services, platforms, and devices rely on computing and technology professionals – who work as data scientists, programmers, or artificial intelligence (AI) experts – to meet their intended goals. How do we train future professionals to have an ethical mindset in their understanding, design, and implementation of algorithms? This was the question that prompted the use of a role-playing case study, which we designed, implemented, and studied in an undergraduate engineering course. We used the Boeing Max 8 flight disaster as the scenario for this case study as it encapsulates how a software algorithm shapes decision-making in a complex scenario. Theoretically, our work is guided by the situated learning paradigm, specifically the need to learn perspectival thinking for decision-making. The ability to make ethical decisions relies to a large extent on the ability of the decision-maker to take context into account – to understand not just the immediate technical need of the work but also larger implications that might even result from unanticipated consequences. Findings from the evaluation of the role-play scenario show that students reported a higher engagement with case study material and a better understanding of the scenario due to taking on a specific role related to the scenario. Analysis of pre-and post-discussion assignments shows a shift in their perspective of the case, further supporting the overall goal of developing a more situated understanding of ethical decision-making.

Using the Boeing Max Air Disaster as A Role-Play Scenario for Teaching Ethical Thinking

Introduction

Algorithms and software now drive decision-making across a range of domains, and the impact of algorithms in the digital future is hard to escape (O'Neil, 2016; Pasquale, 2015). They not only shape individual actions but are increasingly embedded into the structure of our society, making decisions at scale about who gets a loan, who gets hired or fired from a job, who receives social services, who can vote, who goes to prison, for how long, and when they get parole (Brayne, 2017; Buolamwini & Gebru, 2018; Eubanks, 2018). For instance, a software system with COMPAS machine learning-based recidivism risk assessment tool assigns disproportionately higher risk to African-American defendants than to Caucasian defendants (Larson, Mattu, Kirchner & Angwin, 2016). New York City Department of Transportation (NYCDOT) has employed the Manhattan Traffic Model that fails to disaggregate transportation data by gender (Perez, 2019), thereby placing an extra burden on caregivers, predominantly female care workers, who must take multiple buses to navigate the city's periphery. Software systems used for policing that deploy predictive algorithms like Chicago's Strategic Subject List (SSL) disproportionately target African American individuals and neighborhoods for increased policing (Brayne 2017; Ferguson 2017; Karppi 2018; Sheehey 2018). The significance of systems that are engineered by technologists for society requires us to think seriously about preparing workers for a novel and uncertain future guided by software and algorithms (Stevens, Johri & O'Connor, 2014). Specifically, how do we prepare the future workforce to be consistently reflective so that their actions enable a better future with minimal or/and no harm? In other words, how do we help students develop an ethical mindset?

We believe that it is within their academic training that future technologists can be best prepared to develop an ethical mindset and can be equipped to respond to the challenging decisions they will have to make when they enter the workforce. The university is a critical site for this training because future workers will have little time to gain ethical training on the job in fast-paced technological work environments and, further, it might be in the best interest of companies to not provide such training or to 'do no harm' even when individual technologists might otherwise be so inclined. Even when industries are regulated, organizations often spend more resources trying to circumvent those regulations for profit-making than working within the regulations. This is true not just of companies that impact the environment but also new companies such as Uber or Airbnb (Nieuwland & van Melik 2018; Rauch & Schleicher 2015; Rosenblat 2018; Stone 2017). Given that in the digital era, laws and regulations far lag technological development, self-regulation in some form will be essential to our social well-being. It is also within their academic training, early in their professional socialization, that an ethical mindset can be best imparted to engineering students so that they have multiple opportunities to think through these complex issues and so that it becomes a part not just of their thinking but also their identity.

Engineering Ethics Education

The field of engineering ethics has a growing pool of scholarly work over the last decade. Much of the growth in published work can be attributed to maintaining accreditation with professional associations such as ABET (Accreditation Board for Engineering and Technology). A literary study of 26 published engineering ethics interventions in scholarly journals between 2000 and 2015 noted that "The majority of the 26 analyzed articles justified their study by referencing or acknowledging ABET accreditation" (Hess and Fore 2018). The guidelines presented in

Accreditation Criterion 3 address student learning outcomes towards ethical and professional responsibilities, although they do not define a framework or goals to achieve this outcome (ABET pg. 5). Thus, engineering programs have both the opportunity and responsibility to define their own goals towards establishing their ethics curriculum. Colby and Sullivan (2008) found a wide variety of techniques and granularity of engineering ethics content through an analysis of engineering courses.

Among engineering education researchers, there is a consensus that ethics coursework is vital to encouraging an identity of responsible engineering (Herkert 2000; Hess and Fore 2018). One reason for this is the impact of engineering decisions, which can result in catastrophic events (Boisjoly et al., 1989; Birsch & Fielder 1994; Herkert et al. 2020). However, this cannot be the only rationale for teaching engineering ethics. Harris (1996) argued that "Although [tragedies, disasters, and scandals] should be a matter of concern to all of us, and especially to engineering students, who can learn from past mistakes, they are exceptional rather than ordinary occurrences. Most engineers will never be involved in such newsworthy circumstances." However, for engineers in today's world, the context and pace with which engineering decision-making occurs have changed this dynamic. Engineering professionals, even those who are early in their careers, are tasked with applying technology, scientific understanding, and technical breakthroughs to solve, build, and create solutions that have a global impact. In doing so, they are faced with ethical dilemmas that they must bridge in order "to make judgments about what is right and what is wrong" (Heller 2012). These decisions are often not simple, straightforward and are faced with ambivalence towards whatever solution is chosen.

Engineering ethics interventions can take the form of several active learning techniques based on the institution's approach. The types of engineering ethics interventions include broad institutional college-led initiatives such as courses and programs but can also be implemented at a micro view through case studies and role-play scenarios (Hess and Fore 2018). In a study from Prince (2006) regarding role-play instruction for diverse student groups, it was noted that the role-plays were more successful with junior and senior students in engaging critical thought towards engineering ethics as compared to mature, foreign-trained professionals. Role-play scenarios have also been used beyond the field of engineering to enhance ethics instruction. In a study on students in pharmacy law in Saudi Arabia, the participants described role-playing engagement to facilitate a more real-world representative curriculum, which led to a positive learning experience for the students involved (Al-Aqeel, 2013). Overall, studies have begun to provide insight into the diversity of engineering ethics interventions in scholarly work. The complexity, opacity, and ambiguity of engineering decisions make it challenging to produce generalizable principles. Thus, role-playing scenarios and case studies play a critical role in giving students experience with how principles can be applied in these complex, opaque, and ambiguous decision-making moments.

Situated Learning for Perspectival Thinking

The objective of the work we present in this paper was to bring a more nuanced and context-based understanding of how algorithms work in real-world applications, and one of the theories of learning that allow us to design and assess that goal is "situated learning" (Lave & Wenger, 1991; Greeno, 2011; Greeno & Nokes-Malach, 2016, Sawyer & Greeno, 2008; Johri & Olds, 2011; Johri, Olds & O'Connor, 2014). The focus of situated learning on the role of context and on taking

different perspectives guides us in thinking of pedagogical tools that support those activities. Of particular relevance to the proposed work is the idea within the situative perspective that what we learn and how we learn it are closely linked. This situation, or context, determines what people look for, how they get the information, and how they interpret the information. Knowledge arises dynamically and is constantly constructed and/or reinterpreted (Clancey, 2009). Within the situated learning tradition, Greeno and van de Sande (2007) and van de Sande and Greeno (2010, 2012) advance a perspectival view and argue that a person's or group's knowledge and understanding of any conception is "their ability to construct perspectival understandings that are situated in activity and that are organized according to principles that are taken as defining the conception (pg. 14)". This process of developing a perspectival understanding is a continuous process as long as the learner is engaged with the situation. Taking a holistic view, the analogy they provide in their paper is that of a painting or photograph where the focus of a perspective is on the different elements and their relational arrangements. Their formulation of perspectival learning is helpful in designing curriculum that is driven by the case or the situation.

Role-Plays for Engineering Ethics Education

Teaching engineering ethics using role-play activities or simulation has recently gained traction within engineering ethics education. Role-play scenarios are a popular engineering ethics intervention because they allow students to engage in empathy and perspective building exercises in the classroom. This participation is fundamental to early-career engineers, who Loui (2005) argues are likely to face ethical dilemmas that include conflicting viewpoints during their everyday work lives. It is important to recognize that "students come to a course with a variety of backgrounds and developmental stages, different students internalize different ideas" (pg. 383). As there is the certainty that students will face these ethical dilemmas but have different readiness levels, researchers argue that the focus is on readying students to empathize. Hess and Fila (2016) argue that empathy and perspective-taking exercises should be incorporated in the engineering curriculum and are fundamental to holistic engagement efforts. As Doorn and Otto (2013) describe, role-play scenarios serve as an instrument to guide students to engage with, debate, and evaluate decisions from the perspective of different roles. The authors also acknowledge that by making students aware of other perspectives, they better understand pressures and influences that would otherwise have been hidden from view.

Role-playing scenarios (RPSs) promote an active learning environment beyond what is possible in a traditional classroom and encourage students to contextualize the case or scenario they are working on (Shaw 2004; Loui 2000) and engage in sensemaking (Johri, 2009). RPSs provide a collaborative learning pedagogical approach that is effective because 1) collaboration triggers cognitive processes associated with learning, including perspectival thinking (Hmelo-Silver, Chinn, Chan & O'Donnell, 2013); 2) collaborative activity allows learners to strengthen understanding of material they have already learned and repair mental models that maybe fragmented or incomplete (Webb, 2013); and, 3) a cognitive-elaboration approach within collaborative learning requires actively processing information, and aims to elaborate basic information-processing activities such as encoding, activation of schemas, rehearsal, metacognition, perspective, and retrieval (O'Donnell & Hmelo-Silver, 2013; pg. 5).

In terms of how collaboration is structured within RPS, it assists with perspectival understanding and change in perspectives by several mechanisms: First, through the use of prompts or prior

information that is provided to participants to encourage high-level discussion, sharing of perspective, and construction and sharing of explanations; second, through the use of structured controversy whereby the topic presentation of views itself shapes discussion (Easton, 2006; Johnson, Johnson & Smith, 1996); and, finally, through cognitive role specialization where adoption of specific roles requires learners to carry out particular cognitive activities (Webb, 2013, p. 29-30). Role-playing provides an even closer approximation to experience than a case study, which is commonly used in ethics teaching, since "being" a character in a role-play introduces a social component that cannot be achieved with a case discussion (Hertel and Millis, 2002).

Proponents of role-playing argue that it supports learning across a range of areas, including perspective-taking (Pusateri, Halonen, Hill, & McCarthy, 2009), critical thinking (Poling & Hupp, 2009), and communication skills (Nestel & Tierney, 2007). In addition to these skills which are necessary regardless of the domain of learning, role-plays are also useful in teaching and consolidating student knowledge in a given subject area (DeNeve & Heppner, 1997; McCarthy & Anderson, 2000; Poling & Hupp, 2009; Poorman, 2002). In RPSs, students must not only apply their knowledge, but must also negotiate with another character with different information, perspectives, and motivations. As Brummel et al. have argued, this social component is important for students to experience, not only for developing communication and negotiation skills but also for cultivating a greater appreciation of the crucial role that social factors play in resolving an ethical issue. Prior research on the use of RPSs has shown that in their initial reactions to RPS, learners have reported finding the RPS uncomfortable. This discomfort is often due to the experience of "disjuncture" (Jarvis, 1987, pp. 79–84) between the experience of the RPS and the student's previous conceptualizations of a given topic. Studies have shown that these disjunctures can prime students to engage strongly in efforts to develop knowledge and skills to bring "equilibrium" back to their conceptualization (Jarvis, 1987, p. 80).

While researchers have argued that role-play scenarios are a viable means to engage students in engineering ethics topics, there is less evidence available on the extent to which the perspectival thinking among learners can change through role-play scenario participation, in addition to the students reported engagement with case materials.

RQ1: How well do RPSs work for furthering a situated understanding of ethical thinking related to algorithms among learners?
RQ2: What is the nature of perspectival thinking among learners, and how does it change through RPS participation, especially how does adopting different roles shape their thinking, and are students able to adopt a more responsible perspective?

Implementation of the Role-Play Activity
In this section, we discuss the class context in which the role-play activity was included, the role-play scenario, descriptions of the individual roles, and how the activity was conducted.

Course
The role-play was integrated into the curriculum for a university-level class for ethics in information technology. The course is a required core course and fulfills the undergraduate requirements towards various tracks within the IT major. The course is taught through both synchronous and asynchronous delivery in online format. The learning environment is discussion-

focused and uses case studies from different regions around the world. Course material is traditionally presented through a series of videos, readings, quizzes, discussions, and reflection assignments. The topics covered include, among others: global economic history; IT and global development; AI, algorithms, and fairness; and IT and ethics.

For this iteration of the course, role-playing elements were introduced to engage students in an online environment. The course is divided into four modules: 1) Societal Impacts of Technology, 2) Privacy and Surveillance, 3) Data & Algorithms, and 4) Ethics. This study focuses on the Societal Impacts of Technology and introduces Ethical decision-making concepts. The content assigned each week included assigned readings, including both journal and news articles, discussion posts, and homework assignments to scaffold learning, leading up to the goal of creating an individual concept map. Toward the end of the module, the group role-plays activity served as a culminating experience to synthesize all the material covered in the module and allow students to engage with several roles and the perspectives therein. Students were given several resources to allow for preparation for the role-play scenarios. The role-play was outlined in detail, describing the scenario and key individuals associated with each scenario.

Module Schedule	
1	Introduction to module
2	Videos covering module material
3	Readings (scholarly and news articles) covering module material
4	Role-play pre-questions
5	Individual concept map
6	Role-play activity
7	Group concept map
8	Role-play post-questions

Role-Play Scenario details

For this study, the role-play scenario assigned to students was based on recent Boeing 737 Max incidents in late 2018 and early 2019. The incidents involved two catastrophic crashes of the 737 Max passenger jet. As Herkert et al. (2020) describe the scenario:

> "In October 2018 and March 2019, Boeing 737 MAX passenger jets crashed minutes after takeoff; these two accidents claimed nearly 350 lives. After the second incident, all 737 MAX planes were grounded worldwide. The 737 MAX was an updated version of the 737 workhorse that first began flying in the 1960s. The crashes were precipitated by a failure of an Angle of Attack (AOA) sensor and the subsequent activation of new flight control software, the Maneuvering Characteristics Augmentation System (MCAS). The MCAS software was intended to compensate for changes in the size and placement of the engines on the MAX as compared to prior versions of the 737. The existence of the software, designed to prevent a stall due to the reconfiguration of the engines, was not disclosed to pilots until after the first crash. Even after that tragic incident, pilots were not required to undergo simulation training on the 737 MAX."

Many instructors use this incident as a case study for demonstrating real-life ethical dilemmas and incidents driven by software engineering.

Students were given roles of various members of an "Aviation Transportation Investigative Committee" (ATIC), tasked by U.S. Congress to investigate the incident to better understand the lessons that could be learned from the Boeing 737 Max incident with the intent of ensuring future safety and to prevent future air transportation disasters. The members of ATIC were given two open-ended questions: 1) why the incident happened and 2) how it could have been prevented as a preparatory assignment to get the students ready for the role-play activity. They were also tasked with bringing potential recommendations by which Congress could ensure future safety and transparency and rebuild trust among consumers.

Students were given a role to play during the activity. Students were provided with a detailed background and information regarding the alignment of the roles within the activity. The roles for the Boeing scenario (see Appendix) were the following.

Role	Role Title	Role-Play Summary and Alignment
A	Aviation Consultant	A served as an expert witness on similar committees to ATIC; is always keen on expressing the viewpoint of pilots. A is concerned that authority for decisions during flights has shifted from pilots to technology and that decisions about pilot training have been determined by business interests rather than pilots' needs.
B	Professor of Aerospace Engineering	B is an expert on aeroelasticity, specifically nonlinear aeroelasticity flight dynamics of highly flexible wings. B provides insight regarding the change to the wing placement to incorporate the new, bigger, engine on the 737 Max. B does not have all the information needed to judge the decision and is a neutral participant.
C	Software Engineer	C is an expert on human-automation interaction and has designed written software for aerospace applications. C has seen the complexity of interdependent systems firsthand due to having designed software to manage it and to simulate outcomes.
D	Retired FAA Officer	D has knowledge and expertise on the role of Federal Aviation Agency (FAA). D is concerned with salvaging the reputation of the FAA as it has been affected by incident.
E	Boeing Representative	E has worked at Boeing for over 30 years and serves to bring the company's perspective to ATIC. E understands the impact of the incident on the company and wants to engage dialogue on what can be done now.
F	Family Lawyer	F represents passengers and their families. F wants to ensure passengers voices do not get stifled among the "experts" and that companies work towards safety rather than focusing on budgeted consequences.

Data Collection

This study was conducted at a large public university in the United States with undergraduate students in the engineering school who were majoring in technology-related areas, including IT and cybersecurity. The student population is highly diverse and includes many first-generation college students, first-generation immigrant students, Hispanic, and Black students. Data were collected from one semester from two different sections of the same course that covered topics related to technology and society and fulfilled the program's ethics requirements for accreditation purposes. Data were collected initially in the Fall 2020 semester through a written assignment, during the role-play scenario through an approach similar to a focus group (Krueger & Casey, 2014), and through a survey. The student responses were prompted by specific questions posted for them. The Institutional Review Board approved the research study of the institution, and only responses from students who consented to participate were included in the analysis. Fifty-one students took part in the role-play spread across twelve groups. There were 27 participants in one section and 24 participants in the second section. Students reported that they have never participated in a role-play scenario in the past.

Findings

We present findings related to students' reported learning experience and their shift in perspective due to participating in the role-play. There were three general themes among student responses to the role-play activity: students learned much regarding the details and complexity of the case, engaged in scenario-specific research beyond the resources provided, and demonstrated a change in their perspectival thinking.

The first common theme among the responses was that students were aware of the Boeing case, but did not realize the complexity or the different stakeholders. Overall, students reported that they learned new aspects of the Boeing case as a consequence of the role-play discussion. Almost all students either agree (25%) or strongly agree (71%) with this sentiment (Figure 1). Students also reported that their overall understanding of the scenario increased as a result of the role-play (Figure 2). By engaging with all the case materials, students demonstrated a more complete view of the events that led to the disaster and isolated several ethical dilemmas at play.

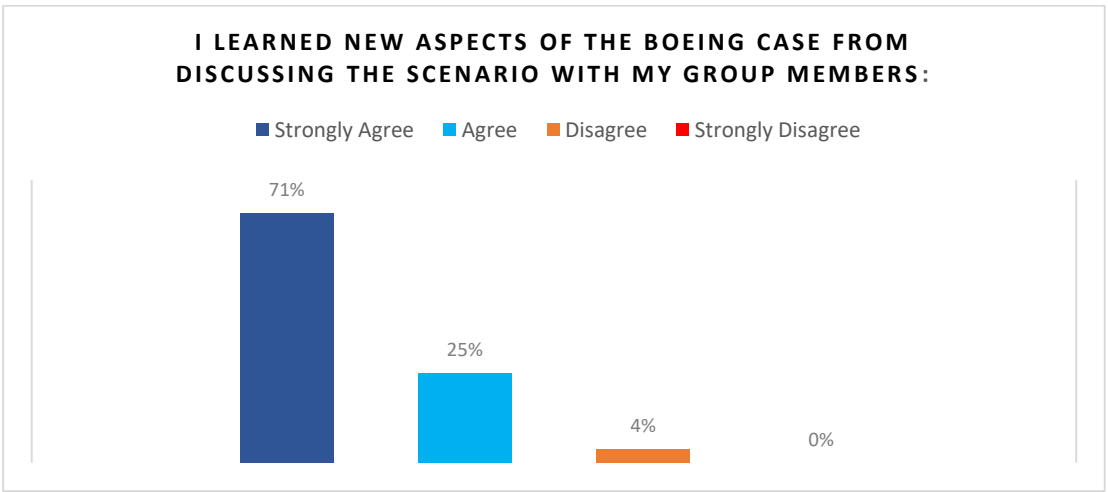


Figure 1: Student responses to the prompt, "I learned new aspects of the Boeing case from discussing the scenario with my group members." (n =51)

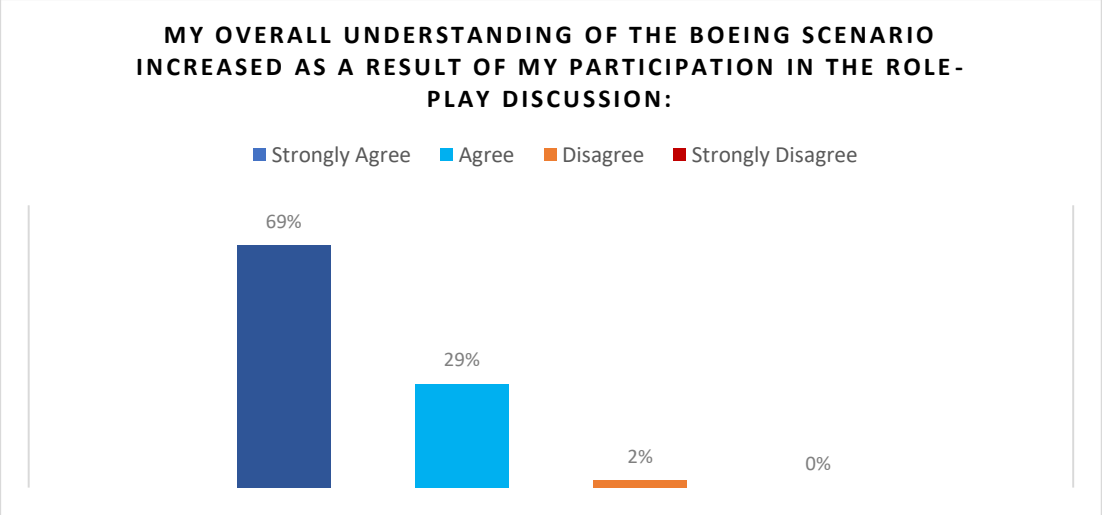


Figure 2: Student responses to the prompt, "My overall understanding of the Boeing scenario increased as a result of my participation in the role-play discussion." (n =51)

Throughout the discussion, different perspectives were expressed, and students reported that although beforehand they held a certain viewpoint, they were able to see things differently after participating in the role-play discussion. Some discourse related to this change can be seen below.

> *My role opened my eyes more to the situation. At first, I thought the MCAS system didn't work, or was badly designed, or implemented. But it was doing exactly what It was meant to do, which I didn't know at first. It seemed that the regulations and guidelines were what let the system down. If the sensor malfunctioned, MCAS would take over and that's what it was programmed to do. So, listening to others talk definitely made me see that perspective. - [Student 1]*

A second theme among the responses was how students engaged with the resources provided and used this to springboard their curiosity to research their own role and the scenario in general. The learning they reported occurred not just through the discussion but as an outcome of preparing for the assigned role. This was partly attributed to the variety of materials used to engage students understanding within the module. Students were provided with academic readings, interviews, reports from different stakeholders, and a list of resources to prepare for the role-pla, along with the role description. Although these items were sufficient to conduct the outcomes of the role-play, each of these items gave direction to additional optional resources that students could use to do additional research. Even though all the information they found may not have been used in the actual discussion, they still reported it as a learning experience. Some discourse from this theme can be seen below.

> *My role's description said something about "human-automation interaction," and at first, I didn't know what that meant. But like I looked it up, and that really got me thinking*

about how much software engineering must work to harmonize the pilot, plane, onboard systems with the software itself. - [Student 2]

During my preparation for my role (aviation/aerodynamics expert). I kind of over-prepared for the role-play… I researched so much about the case because I wanted to be the expert. While in the end we may not have gone it details about thrust or how the wings were shaped, I thought it was very interesting to be able to dive into those aspects of the case because of my role. I wouldn't have thought about it like that without my role. - [Student 3]

Given the centrality of the roles in the discussion, we asked students about their experience with the role and if they would have preferred to play a different role or learned more if they had a different role. Although a majority of the students (~75%) expressed that they were fine with the role they were assigned, a quarter noted a preference or expected learning gains if they had a different role (Figure 3).

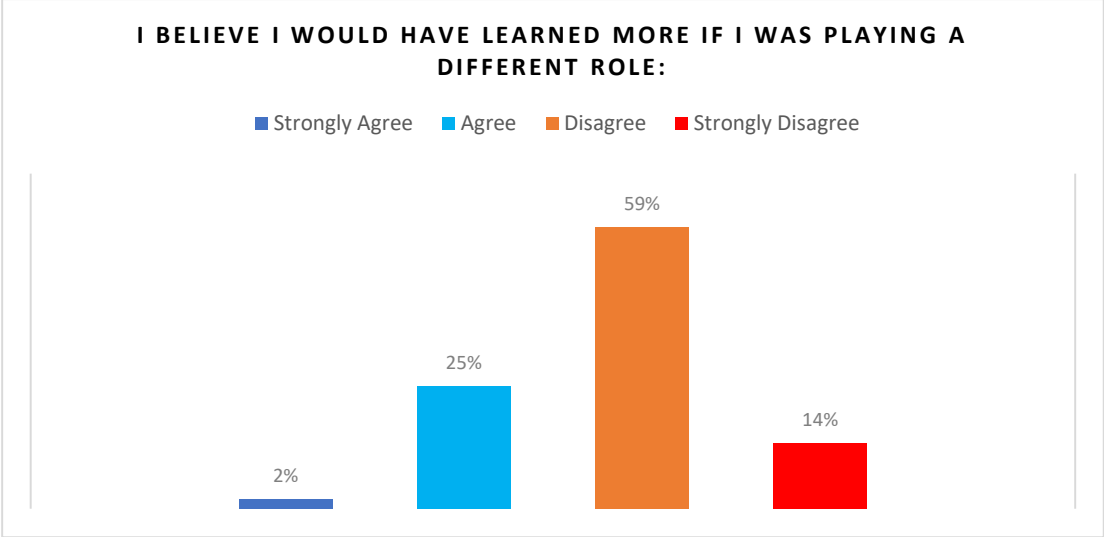


Figure 3: Student responses to the prompt, "I believe I would have learned more if I was playing a different role." (n =51)

We analyzed the data to understand better if this was role-specific, and students reported the least preference for the role of the family lawyer, Role F as that was a challenging perspective for the students to take (Figure 4). Many of the students were also not comfortable playing the role of the Boeing representative which is not surprising as students predominantly reported not agreeing with Boeing's viewpoint in general and the way the situation was handled.

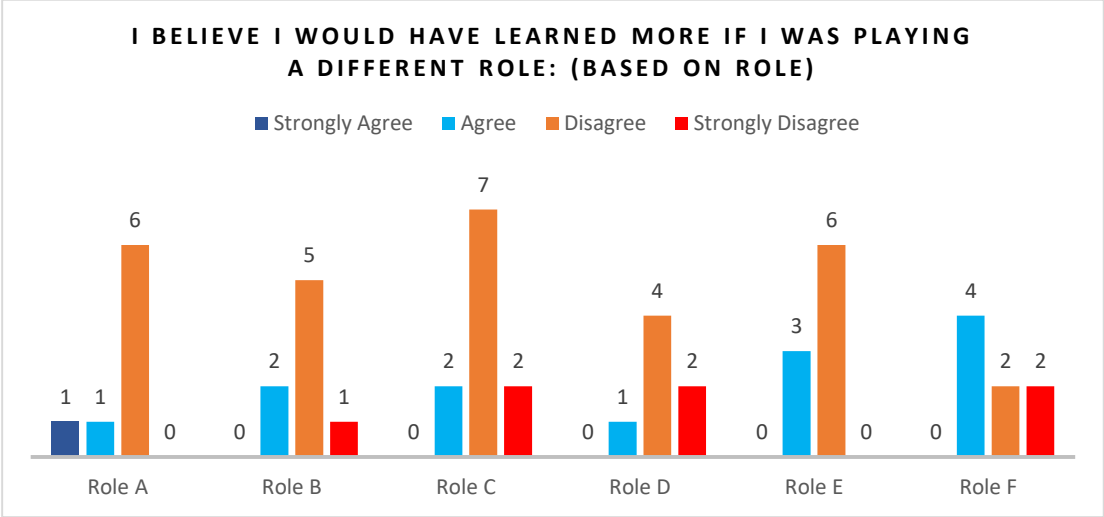


Figure 4: Student responses to the prompt, "I believe I would have learned more if I was playing a different role." based on the role they played. (n =51)

The third theme was the effect of role-playing on the student's perspective. Almost 80% of the students either agree or strongly agree that participating in the role-play activity had changed their perspective (Figure 5). The change in perspective resulted from students expressing different viewpoints and often resulted in a specific role having to change their opinion about what took place. The roles naturally encouraged collaboration and discussion, but since the role-play was focused getting students to form a recommendation, the discussions provided an opportunity for each participant to represent their role. Consensus-focused discussion also made them empathize with other roles and appreciate the overall complexity of the case.

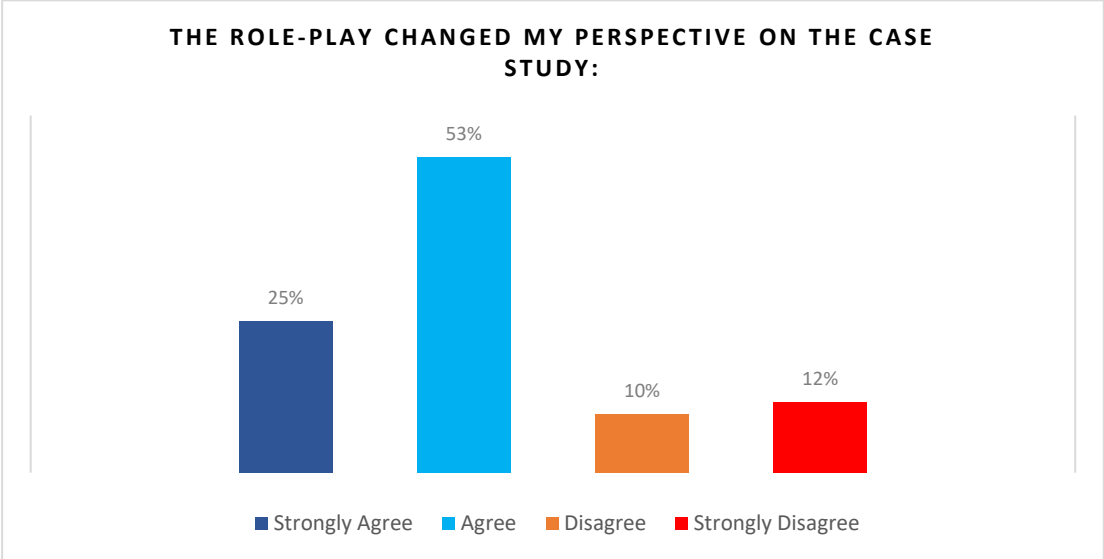


Figure 5: Student responses to the prompt, "The role-play changed my perspective on the case study." (n =51)

Some discourse associated with this theme can be seen below:

> *I think playing the role of the Boeing representative was really tough, but it was really interesting to gain a perspective for the people working at the company. I had to basically disagree with my personal opinion. I know it was tough on for that role, but their conversation showed that there are so many perspectives involved in this case. - [Student 4]*
>
> *Coming into the role-play, I was pointing my finger at the FAA and Boeing, and I still do. But I am more sympathetic to the FAA because of what the retired FAA Officer said during the role-play. It made me realize that the FAA were essentially lied to. I agree there should have been more oversight, but the FAA trusted this company for so long. And, you know, that was wrong, but at the same time, it just made me a little bit more sympathetic towards someone in that position. - [Student 5]*
>
> *I think taking on someone else's role and perspective lets you study more specific information of their area, and when we all come together, we all bring pieces to the puzzle, and we put them together. So, I got a good vision of the scenario and the different perspectives. I also think by doing this, I realized that there is a lot of complexity. It's great to break it off into smaller pieces, and everyone contributes. I enjoyed that. - [Student 6]*
>
> *The role that I played definitely affected my stance and perspective. My role was less technical compared to some of the other roles, but I got to think like a businessperson. I feel like this role was perfect for me because it coincides with my personal perspective on the case. I had a lot of fun with it. - [Student 7]*

Finally, students stated that they did not think that just reading the scenario and writing an essay would have been as effective as actually doing the role-play (Figure 6).

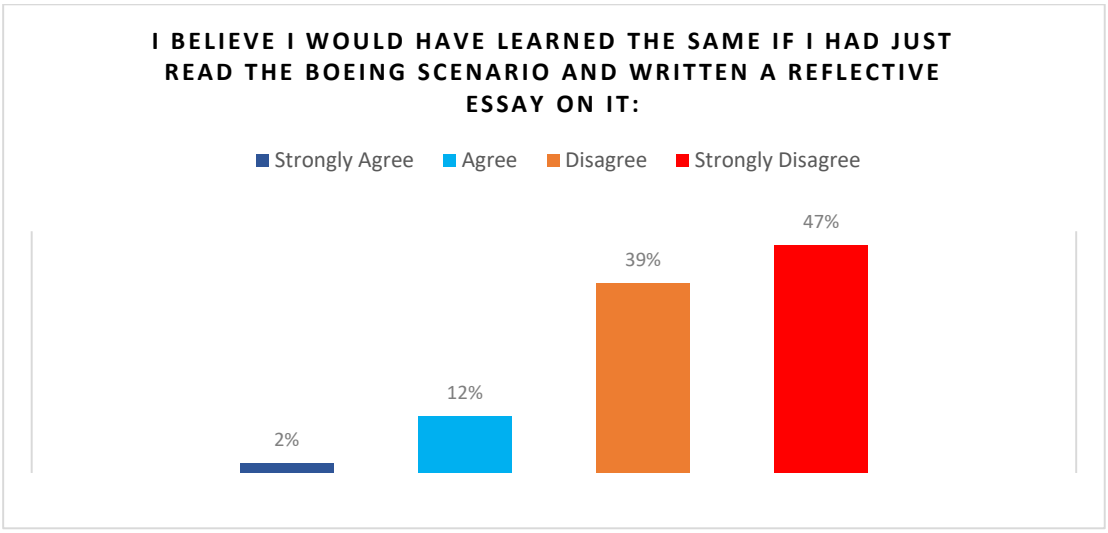


Figure 6: Student responses to the prompt, "I believe I would have learned the same if I had just read the Boeing scenario and written a reflective essay on it." (n =51)

Discussion and Conclusion

In this paper, we present a study of the use of role-play case studies for teaching perspectival thinking related to ethics. We designed and implemented a role-play in an undergraduate course and collected survey and focus group data related to the use of the role-play. We found that the role-play effectively provided students an understanding of how context shapes a situation and how different actors involved with a complex scenario interpret and approach it differently. Through active engagement with peers, schemas, which are the basic cognitive structures and shape perspectives, are enhanced and reshaped, thereby changing how learners view a concept or issue. Peer interaction leads to deeper engagement and also helps learners stay on task. Furthermore, feedback and discussion with peers ensure that learners understand the content they are trying to explain. For the overall success of role-play scenarios in the classroom, students should be both challenged and engaged by the content. There should be an interesting underlying problem that needs to be addressed or a contentious decision that needs to be made. Not all the information should be available in the given information; it should appeal to students' imaginations and encourage them to think beyond what is given.

Acknowledgements

This work is partly supported by U.S. National Science Foundation Award#1937950, "Situated Algorithmic Thinking: Preparing the Future Computing Workforce for Ethical Decision-Making through Interactive Case Studies". Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies. The research study has been approved by the Institutional Review Board at George Mason University.

References

ABET. (n.d.) *Criteria for Accrediting Engineering Programs*. https://www.abet.org/wp-content/uploads/2020/09/EAC-Criteria-2020-2021.pdf

Al-Aqeel, S. A. (2013). Pharmacy Students Feedback on the use of Role-play in Teaching Ethics. *Pharmacy Education*, *13*.

Birsch, D., & Fielder, J. H. 1994. *The Ford Pinto case: A study in applied ethics, business, and technology*. Albany, NY: University of New York Press.

Boisjoly, R. P., Curtis, E. F. & Mellican, E. (1989). Roger Boisjoly and the Challenger Disaster: *The Ethical Dimensions*. *Journal of Business Ethics*, 8(4), 217–230.

Brayne, S. (2017). Big data surveillance: The case of policing. *American Sociological Review*, 82(5), 977–1008.

Brown, K. M. (1994). Using Role Play to Integrate Ethics into the Business Curriculum a Financial Management Example. *Journal of Business Ethics*, 13(2), 105–110.

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Fairness, Accountability and Transparency* (pp. 77-91).

Clancey, W. J. (2009). Scientific Antecedents of Situated Cognition. In Robbins, P. & Aydede, M. (Eds). *The Cambridge Handbook of Situated Cognition*, p. 11-34.

Doorn, N., & Kroesen, J. (2013). Using and Developing Role Plays in Teaching Aimed at Preparing for Social Responsibility. *Science and Engineering Ethics*, 19(4), 1513–1527.

Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York, NY: St. Martin's Press.

Ferguson, A.G. (2017). *The rise of big data policing: Surveillance, race, and the future of law enforcement*. New York, NY: NYU Press.

Greeno, J., Collins, A., & Resnick, L. (1996). Cognition and learning. In R. Calfee & D. Berliner (Eds.), *Handbook of educational psychology* (pp. 15-46). New York: MacMillan.

Greeno, J. G., & Nokes-Malach, T. J. (2016). Some early contributions to the situative perspective on learning and cognition. In M. A. Evans, M J. Packer, and R. K. Sawyer (Eds.), *Reflections on the Learning Sciences* (pp. 59-75). Cambridge University Press. New York, NY.

Greeno, J. G. (2011). A situative perspective on cognition and learning in interaction. In T. Koschmann (Ed.), *Theories of learning and studies of instruction* (pp. 41-72). New York: Springer.

Greeno, J. G. & van de Sande, C. (2007). Perspectival understanding of conceptions and conceptual growth in interaction. *Educational Psychologist*, 42, 9-23

Harris, C.E., Jr., Davis, M., Pritchard, M.S. and Rabins, M.J. (1996), Engineering Ethics: What? Why? How? And When?. *Journal of Engineering Education*, 85: 93-96.

Heller. (2012). Technoethics: The Dilemma of Doing the Right Moral Thing in Technology Applications. *International Journal of Technoethics*, 3(1).

Herkert, J. R. (1997). Collaborative learning in engineering ethics. *Science and Engineering Ethics*, 3, 447-462.

Herkert, J. R. (2000). Engineering ethics education in the USA: Content, pedagogy and curriculum. *European Journal of Engineering Education*, 25 (4), 303 – 313.

Herkert, J., Borenstein, J., & Miller, K. (2020). The Boeing 737 MAX: Lessons for Engineering Ethics. *Science and Engineering Ethics*, 26(6), 2957–2974.

Hertel, J. P., & Millis, B. J. (2002). *Using simulations to promote learning in higher education: An introduction*. Sterling, VA: Stylus Publishing, LLC.

Hess, J. L., & Fila, N. D. (2016). The Development and Growth of Empathy Among Engineering Students. *Proceedings of ASEE*.

Hess, J. L., & Fore, G. (2018). A Systematic Literature Review of US Engineering Ethics Interventions. *Science and Engineering Ethics*, 24(2), 551–583.

Johri, A. (2009). Preparing Engineers for a Global World: Identifying and Teaching Sensemaking and Practice Forming Strategies. *Proceedings of 39th ASEE/IEEE Frontiers in Education Conference*, October 18 - 21, 2009, San Antonio, TX. pp. M2D-1-M2D-6.

Johri, A., Bolds, B. M. & O'Connor, K. (2014). Situative Frameworks for Engineering Learning Research. Johri, A. & Olds, B (Eds). *The Cambridge Handbook of Engineering Education Research*, Cambridge University Press, New York, NY.

Johri, A. & Olds, B. (2011). Situated Engineering Learning: Bridging Engineering Education Research and the Learning Sciences. *Journal of Engineering Education*, 100(1):151-185.

Karppi, T. (2018). "The computer said so": On the ethics, effectiveness, and cultural techniques of predictive policing. *Social Media + Society*.

Kilgour, P., Reynaud, D., Northcote, M. T., & Shields, M. (2015). Role-playing as a tool to facilitate learning, self-reflection and social awareness in teacher education. *International Journal of Innovative Interdisciplinary Research*, 2(4), 8-20.

Krueger, R. A. & Casey, M. (2014). *Focus groups: A practical guide for applied research*. Sage Publications.

Lave, J. & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.

Loui, M.C. (2009). What can students learn in an extended role-play simulation on technology and society? *Bulletin of Science, Technology & Society*, 29(1): 37-47.

Loui, M. C (2005). Ethics and the Development of Professional Identities of Engineering Students. *Journal of Engineering Education*, 94(4), 383–390.

Loui, M. C. (2000). Fieldwork and cooperative learning in professional ethics. *Teaching Philosophy*, 23, 139-156.

Nestel, D., & Tierney, T. (2007). Role-play for medical students learning about communication: Guidelines for maximising benefits. *BMC Medical Education*, 7(3), 1-9.

Nieuwland, S., & van Melik, R. (2018). Regulating Airbnb: how cities deal with perceived negative externalities of short-term rentals. *Current Issues in Tourism*, 1-15.

O'Donnell, A. M. & Hmelo-Silver, C. E. (2013). What is collaborative learning?: An overview. In C. E. Hmelo-Silver, C. A. Chinn, C. K. K. Chan, & A. M. O'Donnell (Eds.). *International Handbook of Collaborative Learning,* (pp. 1-15). New York: Taylor and Francis.

O'Neil, C. (2016, September 01). How algorithms rule our working lives | Cathy O'Neil. Retrieved March 03, 2021, from https://www.theguardian.com/science/2016/sep/01/how-algorithms-rule-our-working-lives

Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, Massachusetts; London, England: Harvard University Press.

Poling, D. A., & Hupp, J. M. (2009). Active learning through role playing: Virtual babies in a child development course. *College Teaching*, 57(4), 221-228. doi:10.1080/87567550903218703

Poorman, P. B. (2002). Biography and role playing: Fostering empathy in Abnormal Psychology. *Teaching of Psychology*, 29(1), 32-36.

Prince, R. (2006). Teaching engineering ethics using role-playing in a culturally diverse student group. *Science and Engineering Ethics*, 12(2), 321–326.

Pusateri, T., Halonen, J. S., Hill, B., & McCarthy, M. (Eds.). (2009). *The assessment cyberguide for learning goals and outcomes*. Washington, D.C. American Psychological Association.

Rauch, D. E., & Schleicher, D. (2015). Like Uber, but for local government law: the future of local regulation of the sharing economy. *Ohio St. LJ*, 76, 901.

Rosenblat, A. (2018). *Uberland: How algorithms are rewriting the rules of work*. Los Angeles and Berkeley, CA: University of California Press.

Stevens, R., Johri, A. & O'Connor, K. (2014). Professional Engineering Work. In Johri, A. & Olds, B. M. (Eds), *Cambridge Handbook of Engineering Education Research*, Cambridge University Press, New York, NY.

Stone, B. (2017). *The upstarts: How Uber, Airbnb, and the killer companies of the new Silicon Valley are changing the world*. New York, NY: Brown, Little and Co.

Shaw, C. (2004). Using role-play scenarios in the IR classroom: An examination of exercises on peacekeeping operations and foreign policy decision making. *International Studies Perspectives.*, *5*(1), 1–22.

Appendix – Scenario

Aviation Transportation Investigative Committee (ATIC) 's Report on the Boeing 737 Max Crisis for U.S. Congress

To better understand the lessons that can be learned from the Boeing Max crisis in order to ensure future safety and to prevent future air transportation disasters the U.S. Congress formed an Aviation Transportation Investigative Committee (ATIC) and instructed it to understand the issues from all perspectives and report back to Congress with recommendations.

Brad Jorgensen and Kathy Schmidt, who had both worked in the aerospace industry at some point in their careers and were now staff members serving members of Congress were designated co-chairs of the committee. For assistance, they had assigned different members of their committee, including external experts, different issues to research and present their findings.

When Brad and Kathy took on this responsibility they thought it would be an easy task as the software was at fault but they soon recognized that the solution to the problem might be more complex. They were surprised to learn that competitive pressure from Airbus in terms of losing market share was one of the major reasons for the new model of the aircraft. They were amazed at how quickly decisions were made to be able to compete including changing the design of the aircraft to accommodate the new engine. As they dug in more into existing news coverage, reports, and testimonies they begun to feel the weight of the problem even more.

They were flying to Seattle to speak with Boeing representatives a few months after they started their work and when boarding the flight in Dulles they noticed that some passengers on their flight asked the desk personnel if the aircraft they were going on was a Boeing 737 Max and the one that had crashed. The desk personnel assured them that all Boeing 737 Max aircrafts were grounded. Trust was a key issue.

The initial time period to achieve this goal was set to be 6 months after the second crash but Brad and Kathy quickly realized that they would need more time to better understand the complex structure that makes up the overall ecosystem for the Boeing 737 Max aircraft. They realized that they would need more input in their report from additional external expert members and they reached out to folks from all different areas of concerns and added them to the committee. They are hoping to look towards the future and come up with guidelines that will be enforced not only through the regulatory agencies but also be made public so that there is more trust in the aircraft and the industry.
Today is the final meeting of the committee and five of the external members are scheduled to present their findings in brief and then discuss what they found to come up with a recommendation for the co-chairs to take to Congress.

The questions posed to the committee are:
1) Why did it happen and how it could have been prevented?
2) How can we ensure future safety and transparency and rebuild trust?

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

Utilizing Topic Modelling To Identify Abusive Comments On YouTube

Shubhanshu Shekhar
Department of Computer Engineering
Delhi Technological University
New Delhi, India
shubhanshushekhar_2k17co338@dtu.ac.in

Akanksha
Department of Computer Engineering
Delhi Technological University
New Delhi, India
akanksha_2k17co34@dtu.ac.in

Aman Saini
Department of Computer Engineering
Delhi Technological University
New Delhi, India
amansaini_2k17co46@dtu.ac.in

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498368

***Abstract*—Online video platforms such as YouTube was once regarded as a haven for entertainment, educational, and promotional purposes. Now they have become a breeding ground for spreading toxic behavior, radicalizing content, and political propaganda. We have used YouTube data API v3 to scrape several data related to youtube videos like videos URL, description, view count, and comment information like commenters, comments, and replies. We investigated some hot topics which are prone to abusive comments. For example, bullies are highly existent in racial, teenage lifestyle and appearance, LGBTQ topics, or targeted mainly towards women and girls. We randomly selected a couple of videos from these YouTube channels and used them to research and analyze this project. We have used LDA (latent Dirichlet allocation) to identify dominant topics in a particular uploader's comment section. We have also used the TextBlob library of python for determining the polarity and subjectivity of comments for each uploader.**

***Index Terms*—YouTube, LDA (Latent Dirichlet Allocation), TextBlob, Abusive Comments , Polarity and Subjectivity**

I. Introduction

The last decade is known as Internet Bloom as in previous ten years, many people had joined the internet and caused massive participation on a platform like Youtube. YouTube is the world's second-largest search engine after Google having over 1.9BN users and over 1,000,000,000 hours of YouTube videos are watching in a day, which is more than Netflix and Facebook videos combined. Hate speech, offensive language, sexual harassment, racism, and other forms of abusive behavior have become very common. That is why we have used youtube for our analysis. The part of our concern is the comment section. The comment section is the section that is accessible to everyone in the world to write anything as some people use it to communicate with other peoples or connect with content creators. Some people misuse this platform and use abusive languages against the audience and the content creator, which is non-appreciated. These may reflect a destructive impact on other people and influence them to use more inappropriate contexts online.

Anyone with a user account can indulge in inappropriate behavior by commenting on the videos, which can often lead to other negative comments and spreads like a disease. Moreover, this toxic behavior is unfortunate and hard to deal with it. Such negativity embarrasses and humiliates subjects. For teenagers and young adults, it can be hard to deal with and can be psychologically jarring. Some abusive cross the line into unlawful or criminal behavior. There are several incidents where young adults and teenagers have committed suicide because of online harassment. Due to the overwhelming audience and large user-base, there must be a check on the video's content and the comments because they are visible to everyone. There is a considerable number of people who make abusive comments. Despite social media efforts to combat abusive behavior online, the problem is still emerging[12]. Now they have joined the arms race with criminals, who are constantly changing tactics to escape these platforms' detection algorithms.

Moreover, only one out of ten people will inform their parents or trusted adults about their abuse. It increases the child's risk of anxiety, low self-esteem, depression and can even lead to drugs and alcohol, and insomnia.

The major challenge is that the cyberbullying attack can come at any point of the day. As most teens have their smartphone and 24 hrs access to the internet so they can bully or are being bullied at any time virtual means plus it is very challenging to defend oneself from cyberbullying as we live in a digital world.

One survey has also suggested that 19% of the teens have reported that someone has written or posted embarrassing things on their social network, and 71% of the survey participant do not feel like social platforms are doing enough to fight the problem. Also, 42% of LGBT youth have experienced cyberbullying, while 35% have received online threats. In contrast, 58% have been a victim of hate speech at least once. Hence, we have chosen YouTube mainly because it is prevalent among people. It provides a fertile environment for these abusive behaviors, as most of the videos have a comment section open in them[13]. Some of the questions we will answer in this project are the topics that people are most abusive about and how people are exploiting, that is, what is the worst form of abuse. We present the below contribution in this paper.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:00:31 UTC from IEEE Xplore. Restrictions apply.

1. Exploratory Data Analysis

1.1. Top words that are highly correlated with the uploader and are not appropriate 1.2. Formation of the word cloud to visualise the data and getting insights from it 1.3. Graphically visualising the profanity of comments

2. Topic Modelling using LDA (Latent Dirichlet allocation) , i.e., looking at words in a text and determining what topic is the text all about. Here, we will try to find topics using LDA (Latent Dirichlet Allocation), and from the topics that people talk to each other in the comment we will Figure out which uploader is more prone to abuse.

3. Subjectivity and polarity i.e, we will use inbuilt TextBloblibrary of python and determine the subjectivity and polarity of the comments. By plotting the same, we can get the insights of a positivity - negativity, and fact-options of each uploader's comments. We can then compare the abusiveness to each YouTuber's comment.

II. RELATED WORK

There is no dataset available online for abusive youtube content. Therefore, to do any analysis, we need to extract our data, create our dataset, and research our own accord. Thus there could be analysis on the dataset, which is done through manually labeling the data. Researchers do the work on working on their dataset and doing separate studies.

Previous works on this topic include Inoka Amarasekara and Will J Grant showed the gender gap in Youtube channels[1]. Chen, H., Mckeever, S., Delany, S. J. worked on detecting the Abusive Content of social media by harnessing text mining power [2]. Fox, J., Tang, W. Y. and Whittaker E. Kowalski R.-M who showed the statistics about sexual harassment to women on the social platform[3][5]. Obadimu A., Mead E., Hussain M.N., Agarwal N. identified different toxicity types in the Youtube Video Comments[4].

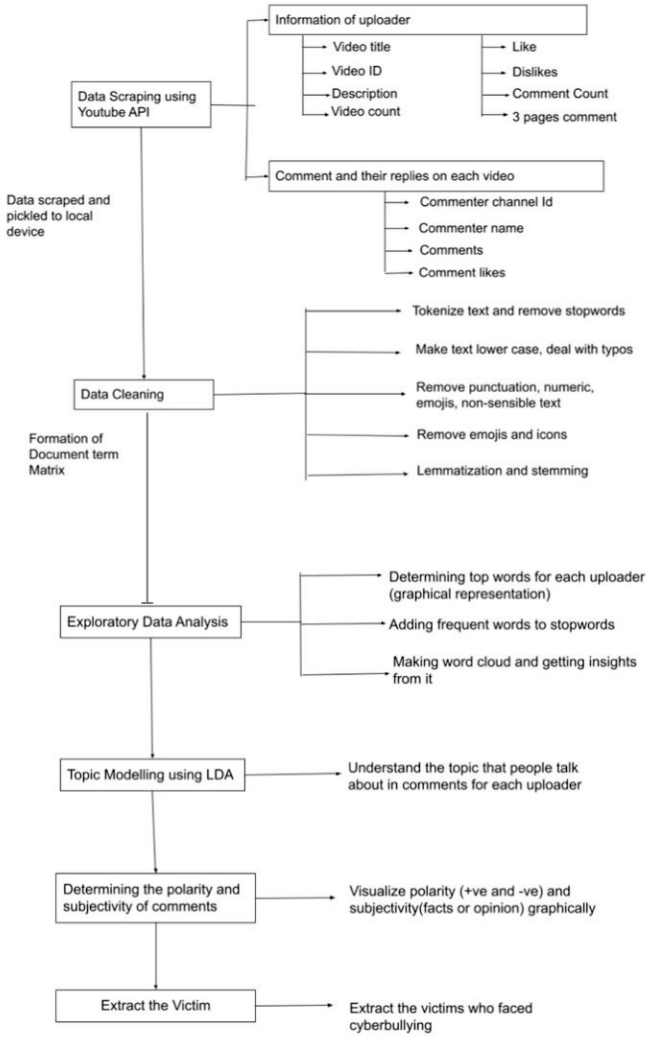
By referencing the above research and their methods, we developed an analysis to Find which Youtuber had the most abusive comments. Thus it needs to report this information to youtube and turn off the comment section for the good well-being of Youtuber and the people who use Youtube.

III. METHODOLOGY

A. Data collection

We have investigated some hot topics which are prone to bullying, for example, teenage lifestyle and appearance, LGBTQ topics, or cyberbullying is targeted mainly towards women and girls. We have also chosen some of the YouTube celebrities who were bullied or cyberbullied at some point in their life and randomly selected a couple of videos from their YouTube channel to research and analyze this project.

We have used YouTube Data API v3 to scrape the metadata of these handpicked videos that is video URLs, description, views and likes statistics, comment information like commenter's name, comments, comment likes, And their replies. We have also analyzed the commentator's and their comments by scrapping all three-page comments per video as a single text. For analyzing what kind of comment each uploader is getting and saving comments and their replies in a separate CSV File for each video.



1. Flow Chart

B. Data pre-processing

The phrase garbage in garbage out is also applicable to the machine learning projects out there. Data preprocessing is an essential step before any NLP procedure. The data scraping part is very loosely controlled and further results in larges chunks of the data out of which only some amount of data is of our need. Hence we need to clean that. Here we combine multiple videos into one by concatenating all the comments into one for a particular user. We are interested in finding what kind of comments a particular uploader gets. After that, we

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:00:31 UTC from IEEE Xplore. Restrictions apply.

applied traditional text cleaning methods, i.e., making all the text lower case, removing punctuation, removing numeric values, removing non -sensical text, tokenize text, and removing stop words and emoticons, and also dealing with typos. Here we are skipping lemmatization and stemming. It is because they change the spelling and map word to their base word, moreover. In this project, we are looking for abuse words too. We do not want such words to be changed. Our goal is to make a Document term matrix (DOM), i.e., uploader name along the row side and number of words along the column side. Finally, we have eight rows and 7143 columns as an output of this step.

C. Exploratory Data Analysis

EDA is used to understand better the main features of data, variable, and relationship that holds them, identifying the crucial variable for us, which is not directly visible.Exploratory data analysis majorly uses statistical graphs and many other data visualizing methods. It tells us beyond what the actual model can tell us, i.e., it gives us the fundamental insight into the data and gives us an intuition of what model to use for machine learning. Here we AIM to explore our data set and find the hard facts about the comments, such as top words for all the videos, top 30 words for each person.



Fig 2. Word Cloud

We will add frequent words to the stop words as if the words are coming in all the unloaders, so it does not make any sense to have it now as it will not help us to differentiate between the videos, hence include that words also to the stop words and remove then while preprocessing. We also here view the graphical relationship between several mild and swear words. Moreover, we have also made the word cloud to analyze the most common words for each uploader.Graphical analysis is done to identify a number of unique words in the comments section for each uploader, i.e., a frequency graph of unique words

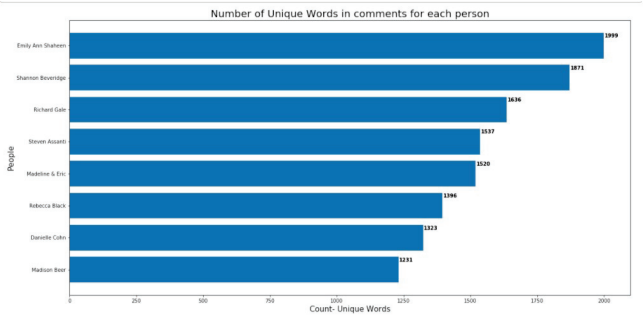


Fig 3. Unique word graph

D. Topic Modelling

Topic Modelling is a statical modelling technique use for discovering the abstract topic from a document. This is an unsupervised classiZication between documents. We can also say that it is analogous to clustering in numerical data. LDA (latent Dirichlet allocation) is the most popular topic modelling technique among all the topic modeling ways. Hence we are using it. Inbuilt python modules are used for this purpose with the parameters of num_topics=4 and passes=20. We have applied LDA on a large chunk of comments proposed before. Here we will understand the topics, i.e., topics for each uploader's comments, and after that, analyze which uploader is more prone to abuse or cyberbullying.

E. Sentiment Analysis

Sentiment analysis is a process of detecting positive, negative, or neutral sentiments in a text. It is mainly used on textual data. Hence we are using it in our project for analytics purposes. Since everyone expresses their feelings and thoughts more openly than ever, analyzing the sentiments had become a significant part of any NLP pipeline.
Using sentiment analysis, we not only focus on polarity, i.e., positivity or negativity, but also on urgency, emotions (happy, sad, low), and even intentions. We used the TextBlob library of python for the sentiment analysis and applied the methodologies in each of these subjects' comments, and analyze what they could mean. We will find polarity and Subjectivity of comments for all the uploaders to determine each uploader's sentiments. Polarity tells us how much positive or negative connotation a sentence or group of the sentence has. It ranges between -1 to 1, where 1 is a super positive sentence while -1 otherwise, and Subjectivity tells us if a sentence states a fact or opinion. It ranges from 0 to 1. 0 is where a sentence or group of a sentence is factual while one is highly opinionated.

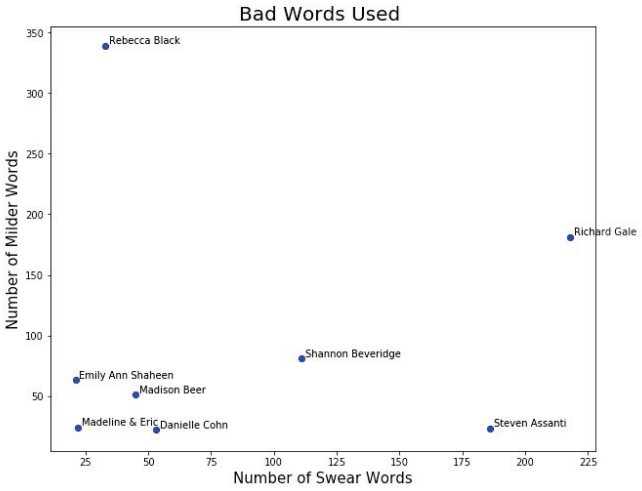
IV. RESULTS AND DISCUSSION

1) In the section of Exploratory data analytics, a graph is made between the number of mild words and the number of abusive words for the understanding purpose about the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:00:31 UTC from IEEE Xplore. Restrictions apply.

frequency of mild and abusive words for each uploader . We can get the following insights from the graph

- Highest Number of Swear Words: Richard Gale
- Second Highest Number of Swear Words: Steven Assanti
- Highest Number of Milder Abuse Words: Rebecca Black
- Least swear words: Emily Ann Shaheen , Madeline Eric
- Least milder abuse words: Madeline Eri



From the chart above, we can confirm that videos on Steven Assanti and Richard Gale has more abuse words. Even from the word cloud we saw that they have words that hurts someone psychologically.

We have used LDA for the topic modelling and the results of topic modelling are as follows

SN.	Topics	Uploader name
1	song/music/voice, Friday/ party/weekend	Rebecca Black
2	beautiful/perfect/amazing song/music/voice	Madison Beer
3	gay/lesbian, sex, kid, fat, victim, b**ch	Shannon Beveridge, Steven Assanti, Richard Gale
4	arab, gold, girls, family, beautiful, pregnant	Emily Ann Shaheen, Madeline & Eric

Table 1. Topic corresponding to each uploader comments

From topics people talk about in the comments for each of the uploaders, we can see that yet Shannon Beverage, Steven Assanti, Richard Gale are prone to harsh comments than any other people on the list. As the number of swear words is much more in the comments of these YouTuber.

Sentiment analyses are done on each uploader's comments to determine the subjectivity and polarity of the comments. Sentimental analysis to get the insights from each uploader's comments, i.e., to determine how positive or negative a comment or how much opinionated each uploader's comments are? i.e., Subjective in a sentence generally refer to feeling, emotions or opinions and judgment, whereas objective refers to hard facts. Polarity is how positive or negative a comment is, i.e., it ranges from negative 1 to positive one, and comments can be positive, negative, or natural depending on the sign of value.

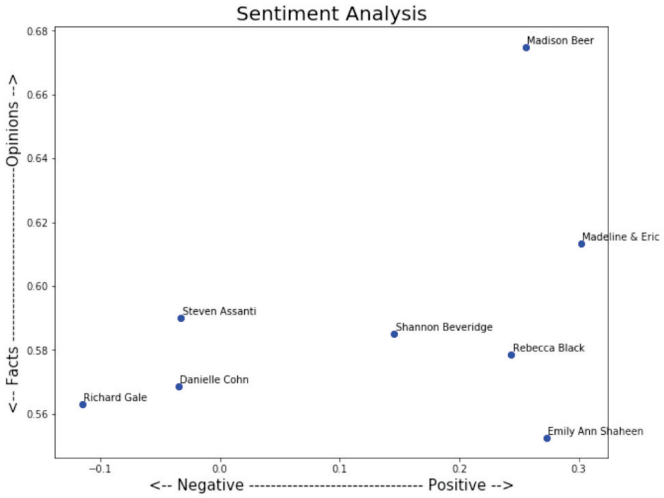


Fig 4. Sentimental Analysis

From the chart we can infer that:

- Madison Beer gets more opinionated positive comments. She is a singer, so it is obvious for her to getopinions at her singing.
- Emily Ann gets positive but factual comments. She gets comments mostly about her Arabic heritage. It is good to know that she does not get much xenophobic comments.
- Madeline and Eric get mostly positive and fairly opinionated comments. They are independent teen parents blogging their livelihood. Most people find it commendable.
- Rebecca Black who used to be heavily bullied for her Friday song now receives factual and positive comments. Though, most of the words in comments of her new video still refer to the 'Friday' song, she is admired for her persistence in singing career.
- Steven Assanti and Danielle Cohn gets more negative comments than above individuals. Steven gets more opinionated comments than Dani. The reason could be he is obese and bit rebellious towards comments on his obesity which generates worse comments. Danielle on other hand is a teenager who displays her odd choices of life in social media.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:00:31 UTC from IEEE Xplore. Restrictions apply.

- Richard Gale videos get the worst comments and also contain factual comments. He was a bully who started a fight with his classmate. However, it is disheartening to see a minor being commented on harshly on social media. There should be some moral policing around this.

V. Conclusion

Looking at all previous comparisons, we find that Richard Gale and Steven Assanti are worse victims of cyber- bullying. Since Richard was/is a minor, we do not want him to take the wrong steps because of such a negative impact. Hence we would contact YouTube administrators or cyber police to ban comments on his videos as the first measure. Moreover, some of the hard-hitting question we can get from here are how to prevent this type of abusive behavior, can mocking will also be considered in the bullying, and do cyberbullying victims, and perpetrators Zit any stereotypical profile or not .

References

[1] Amarasekara I, Grant WJ. Exploring the YouTube science communication gender gap: A sentiment analysis. Public Understanding of Science. 2019;28(1):68-84. doi:10.1177/0963662518786654

[2] Chen, H., Mckeever, S., Delany, S. J. (2016). Harnessing the Power of Text Mining for the Detection of Abusive Content in Social Media. Advances in Computational Intelligence Systems, 187-205. doi:10.1007/978-3-319-46562-3_12

[3] Fox, J., Tang, W. Y. (2015). Women's experiences with general and sexual harassment in online video games: Rumination, organizational responsive- ness, withdrawal, and coping strategies. New Media Society, 19(8), 1290– 1307. https://doi.org/10.1177/1461444816635778

[4] Obadimu A., Mead E., Hussain M.N., Agarwal N. (2019) Identifying Toxicity Within YouTube Video Comment. In: Thomson R., Bisgin H., Dancy C., Hyder A. (eds) Social, Cultural, and Behavioral Modeling. SBP-BRiMS 2019. Lecture Notes in Computer Science, vol 11549. Springer, Cham. https://doi.org/ 10.1007/978-3-030-21741-9_22

[5] Whittaker E. Kowalski R.-M. (2015). "Cyber-bullying via Social Media." Journal of School Violence, Vol. 14, No.1. DOI: http://dx.doi.org/10.1080/15388220.2014.949377 Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[6] Thomas Davidson, Debasmita Bhattacharya, Ingmar Weber. (2019). Racial Bias in Hate Speech and Abusive Language Detection Datasets.

[7] Lange, P. G. (2014). Commenting on YouTube rants: Perceptions of inap- propriateness or civic engagement? Journal of Pragmatics, 73, 53–65. doi:10.1016/j.pragma.2014.07.004

[8] R. Kaushal, S. Saha, P. Bajaj and P. Kumaraguru, "KidsTube: Detection, characterization and analysis of child unsafe content promoters on YouTube," 2016 14th Annual Conference on Privacy, Security and Trust (PST), Auckland, 2016, pp. 157-164, doi: 10.1109/PST.2016.7906950.

[9] Kapadia, S. (2020, December 29). Topic Modeling in Python: Latent Dirichlet Allocation (LDA). Medium. https://towardsdatascience.com/end-to-end-topic-modeling-in-python-latent-dirichlet-allocation-lda-35ce4ed6b3e0

[10] Obadimu, A., Mead, E., Hussain, M. N., Agarwal, N. (2019). Identifying Toxicity Within YouTube Video Comment. Lecture Notes in Computer Science, 214–223. doi:10.1007/978-3-030-21741-9_22

[11] Obadimu, A., Mead, E., Hussain, M. N., Agarwal, N. (2019, July). Identifying Toxicity Within YouTube Video Comment. In International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation (pp. 214-223). Springer, Cham.

[12] Hosseini, H., Kannan, S., Zhang, B., Poovendran, R. (2017). Deceiving google's perspective api built for detecting toxic comments. arXiv preprint arXiv:1702.08138.

[13] Gr¨ondahl, T., Pajola, L., Juuti, M., Conti, M., Asokan, N. (2018, January). All You Need is" Love" Evading Hate Speech Detection. In Proceedings of the 11th ACM Workshop on Artificial Intelligence and Security (pp. 2-12).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:00:31 UTC from IEEE Xplore. Restrictions apply.

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

Very Short-Term Photovoltaic (PV) Power Forecasting Using Deep Learning (LSTMs)

Razzak Ahmad
Electrical Engineering Department
Delhi Technological University
Delhi, India
razzakahmad_2k17ee169@dtu.ac.in

Pratyush
Electrical Engineering Department
Delhi Technological University
Delhi, India
pratyush_2k16ee94@dtu.ac.in

Rajesh Kumar
Electrical Engineering Department
Delhi Technological University
Delhi, India
rajeshkumar@dtu.ac.in

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498536

***Abstract* - Transition from carbon-intensive sources to renewable energy sources is of paramount importance and will curb effects of greenhouse gas emissions. India has witnessed rapid expansion in the solar power sector. Prime utilization of solar power calls for a reliable prediction model. Precise real time data of PV systems is the need of the hour. Market today requires the use of very short-term forecasts. This study aims to present a multivariate approach to forecast solar power generation of PV systems on a very short-term basis using LSTM, a deep-learning algorithm. System performance and capabilities are validated on the basis of a standard metric, Mean Absolute Error (MAE). The model incorporates other features of Machine Learning like hyper parameter tuning and optimum feature selection which further refines the accuracy.**

Keywords – Long short-term memory (LSTM), Photovoltaic forecasting, Recurrent neural network, Time series prediction

I. INTRODUCTION

The International Energy Agency (IEA) has projected a growth of 110 GW per year in global net solar PV capacity additions by 2023. With India's annual PV additions almost doubling in 2021 compared to 2020, it shall play a crucial role in the renewables upswing [1]. Corresponding the world's concern to tackle climate change, Government of India along with the ISA, recently came up with the concept of One Sun One World One Grid (OSOWOG). Implementation of this global electricity ecosystem will put India in the pole position to ensure fulfillment of low-cost energy demand of the world [2]. Solar power has already replaced non-conventional sources of energy as a cheaper alternative for electricity production in many countries. In anticipation, by 2024 total installed renewable energy capacity (solar and wind) will exceed that of coal [1]. Growing levels of energy consumption have led to a drastic fall in the availability of conventional resources of energy resulting in a massive energy crisis all across the globe. Moreover, greenhouse gas emissions (GHGs), are the cause of environmental issues like global warming. Trending as a viable means of power production, solar power has gained ground for both its environmental and economic benefits [3].

However, like many other green-energy sources, photovoltaic power faces challenges of volatility and instability for its dependency upon meteorological conditions. Today, when integration of renewables into the grid is at a faster pace than ever, uncertainty in the PV outputs may deteriorate security of the power grid leading to soaring operational costs and more importantly severely affect the quality of life of consumers. Hence, development of an appropriate forecasting tool has captivated both the research community and other stakeholders of the energy market. By mitigating the intermittent nature, an accurate prediction technology can improve flexibility of the power generation system and help maintain synergy between the demand and consumer side of PV power market [4]. This study revolves around very short-term forecasting where the forecast period does not exceed more than a few hours. This prediction horizon is advantageous to power plant and grid operators, and the ever-increasing consumers of electricity [3]. Short- term predictions complement electricity price forecasting and play a pivotal role in solving issues related to decision- making in the energy market which may include economical dispatch planning and unit commitment [5].

Reference [6] shows that based on the methodical approach used, forecasting methods can be commonly classified into:

(i) physical models, (ii) statistical approaches, and (iii) hybrid approaches, a combination of the above methods. A physical relationship between input meteorological characteristics and required output solar power make a physical model which finds application in models devoid of any form of historical input data [7], [8]. Reference [9] shows that historical time- series data which is often stochastic in nature forms the basis of statistical approaches categorized as: i) classical techniques like time series [10], [11]; or ii) advanced which includes powerful machine-learning methods [12], [13]; and lastly iii) deep-learning approaches have been on the rise lately, specifically those based on neural networks, as far as short-term forecasting is concerned [14]. Varying statistical measures show hybrid methods' viability in forecasting on grounds of simplicity, affordability and accuracy [15], [16]. Increasing size, diversity and inter-connections in dataset lead to more noise and complexity. This causes an influx of issues, like difficulty in scaling of data for various traditional statistical, ML and hybrid methods. Consequently, deep learning algorithms have been the anchor of various temporal forecasting assignments. Hence, being an active field of research, DL architectures have been heavily exercised in regenerative power forecasting especially those using short- term horizons [17]–[19].

Inspired by the forecasting strength of deep-learning in aforementioned domains, the proposed research work analyzes the effectiveness of LSTM-RNN in forecasting photovoltaic power generated from a PV setup located in Berkeley, California. References [17], [19] manifest the ability of LSTM (long short-term memory) in learning time- series dependencies. Empirically, it is seen that the choice of appropriate input features is significant than the size or dimensionality of dataset. Hence, one of the crucial aspects addressed by this paper is the variable selection technique which finalizes the set of most influential weather attributes needed for power prediction [20]. As of now, feature

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:59:05 UTC from IEEE Xplore. Restrictions apply.

engineering using an ensemble learning method has only been sparsely resorted to in a deep-learning model utilized for solar power forecasting. Here feature selection has been executed using Random Forest regression. Additionally, we recommend to study the influence of hyper parameter tuning in optimizing the network performance which has not been investigated in other DL-based literature [17]-[19].

II. DEEP LEARNING

High levels of accuracy along with algorithm adaptability and scalability is the desire of most researchers working in domains requiring the use of machine learning. Statistical as well as classical ML methods have succumbed to problems arising out of colossal, unlabeled and unstructured datasets. Such shortcomings have been perfectly eased out by deep learning which does not require labeling or structuring of data. Concepts such as feature engineering and hyper parameter tuning have further helped reduce errors improving the accuracy of DL models.

Predominantly, recurrent neural networks (RNNs) underpin the whole field of deep learning. Recurrent linking between neurons in the network help in learning and recognizing complex data patterns. Hierarchical extraction of information from lower-to-higher level results in the transfer of data between successive input layers from the previous output layer. This complex mapping between the input-output structures implies that RNNs are largely influenced by the working of a human nervous system [21]. It is explicated that feedback connections among the hidden units allow recurrent neural networks to memorize time-dependent behavior of the datasets [22]. A looping mechanism in the networks mimicking a high-way enables modelling of sequential data needed for addressing time-dependent forecasts.

Nonetheless, traditional RNNs seem to have a downside in the sense that details from the earlier time-steps vanish from the memory and only those from the recent ones are preserved. Vanishing gradient problem caused by the nature of back-propagation algorithm doesn't contribute to ample learning. In consequence to the gradient explosion problem, RNNs fail to learn long-term dependencies when dealing with large-scale datasets. Elimination of this built-in problem prevailing in this neural network arrangement was the principal reason that resulted in the development of a more specialized version known as long-short-term-memory (LSTM) [23].

III. LONG SHORT-TERM MEMORY

Formulated in 1997 by the German duo of Hochreiter and Schmidhuber [24], LSTM with its cascaded formation, exhibits competency by memorizing long-term dependencies. The network memory is well regulated via the opening and shutting of a three-gate system which permits information addition, modification or removal from the cell- state [14]. These gate structures known as the input gate, forget gate and output gate alleviate the issue of short-term memory and enable palpable utilization of multivariate influences in LSTM architectures [25]. Fundamentally, each gate involves a unique tensor operation and there are four such sub-operations during the training of a LSTM cell. Control gates generate their response as input data undergoes logical operation and nonlinear transformation because of activation functions, namely the sigmoid (σ) and the hyperbolic tangent (tanh). Encounter with a high activation directs input gate to add input data to the cell state and an output gate to release information to the next cell. Comparatively less-important data is discarded from the cell state when forget gate notes a high activation [14].

The process starts with concatenation of current input Xt at time step t and ht-1 (previous input vector) to form the input vector. The latter denotes the hidden state at t-1 (previous time step). Ut in (4) refers to the new memory cell candidate value (also known as update value) and the current state of memory cell is represented by Ct as shown in (5). The control flow in a LSTM block can be mathematically formulated by the following equations:

$$i_t = \sigma(W_i \times [C_{t-1}, X_t, h_{t-1}] + b_i) \quad (1)$$

$$f_t = \sigma\left(W_f \times [C_t, X_t, h_{t-1}] + b_f\right) \quad (2)$$

$$o_t = \sigma(W_o \times [C_{t-1}, X_t, h_{t-1}] + b_o) \quad (3)$$

$$U_t = tanh(W_u \times [C_{t-1}, X_t, h_{t-1}] + b_u) \quad (4)$$

$$C_t = f_t \circ C_{t-1} + i_t \circ U_t \quad (5)$$

$$h_t = o_t \circ tanh(C_t) \quad (6)$$

where input gate, forget gate, output gate and input modulation gate are represented by the subscripts i, f, o and u respectively. W and b denote the weight and bias parameters.

The input gate consists of two layers called input gate layer and update gate layer. Considering value of the input vector, this gate determines new information that has to be passed to the cell state. As the name suggests, the second layer of this gate is responsible for updating the current cell state value by adding new candidate values. As seen in (2), the unimportant prior memory state is assessed and then removed by the forget gate. In an update process, the current cell state Ct is fetched from its previous value as demonstrated in (5). Finally, output gate defines portions of the internal cell state that will be used for the hidden state, which is eventually used for prediction. From (6), on passing through the tanh activation function, the vector gives the new hidden state ht which represents output of LSTM at the next moment t. In a solar power forecasting model employing LSTM, it would represent the power output at next moment. Hadamard product (element-wise vector multiplication) is demonstrated by the operator '∘'.

IV. PROPOSED RESEARCH FRAMEWORK

A. Dataset and Framework

Long short-term memory (LSTM) is experimented on data from a solar PV power setup in the city of Berkeley, California. Data is publicly available and can be downloaded from the open repository of Kaggle portal [26]. Compiled with their respective abbreviations in table 1, are seven meteorological features extracted from the dataset using Python 3.6. A year-long monthly data from 1 September 2008 (1:00:00) to 31 August 2009 (22:00:00) recorded on a three-hour interval is used for training and testing the DL algorithm covering 2920 samples in the entire period. Fig. 1 graphically validates the stochastic nature of solar PV power production.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:59:05 UTC from IEEE Xplore. Restrictions apply.

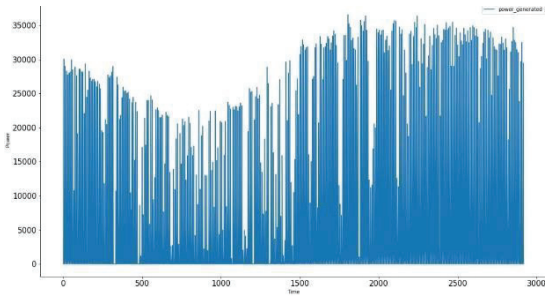


Fig. 1. Solar Power Generation at PV Facility in Berkeley

Fig. 2 presents the framework of the proposed forecasting methodology showing the four main steps described in detail in the subsequent sections.

TABLE I. SUMMARY OF METEOROLOGICAL PARAMETERS

Meteorological Parameter	Abbreviation
Daylight	is_daylight
Distance to Solar Noon	dist_to_solar_noon
Cloud Cover	cloud_cover
Visibility	visibility
Relative Humidity	rel_humidity
Average Wind Speed Period	avg_ws_period
Average Barometric Pressure Period	avg_bp_period

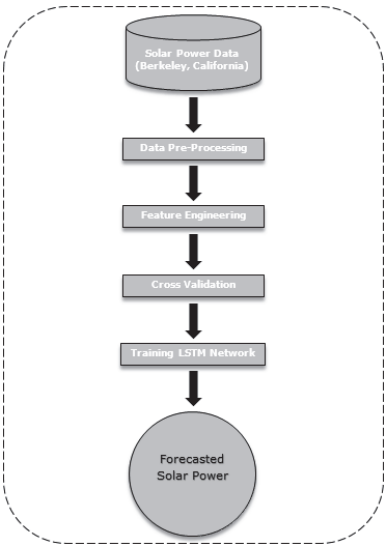


Fig. 2. Steps Followed in the Forecasting Process

B. Performance Metric

Power forecasting being a regression problem asks for a performance metric which can help gauge the forecasting accuracy. Mean absolute error (MAE), being simple in its interpretation, is offered in this research paper with the purpose of representing the various prediction error values in the succeeding sections.

$$MAE = \frac{1}{n}\sum_{k=1}^{n} |e_k - \hat{e}_k| \quad (7)$$

where e_k represents the actual solar power and $\hat{e}_k$ represents its predicted quantity.

C. Data Pre-processing

A poor quality input data may adversely affect the accuracy of any forecasting model. Hence, there is a need of data pre- processing to rectify the issues related to missing data or quality of data. First and foremost, all date and time related data (year, month, day and first hour) in the dataset arranged in separate columns are grouped into a single column named date_time. Secondly, using preceding and subsequent sample points, missing values in the raw dataset have been interpolated. Finally, data is completely processed and stabilized using the MinMaxScaler( ) function of Python scripts which normalizes data in the [0,1] range.

D. Feature Engineering

The right choice of input features is non-trivial for improving accuracy of a forecasting model and may otherwise cause difficulty in resolving complexity of the deep-learning algorithm. Experiential approach to LSTM- based forecasting has proven that enlarging dimensionality of the dataset does not necessarily translate to greater precision of the suggested forecasting approach. This experiment makes use of Random Forest regression variable selection technique. Consisting of multiple decision trees, the working rule of this ensemble method is the recursive removal of irrelevant parameters until the best ones are found.

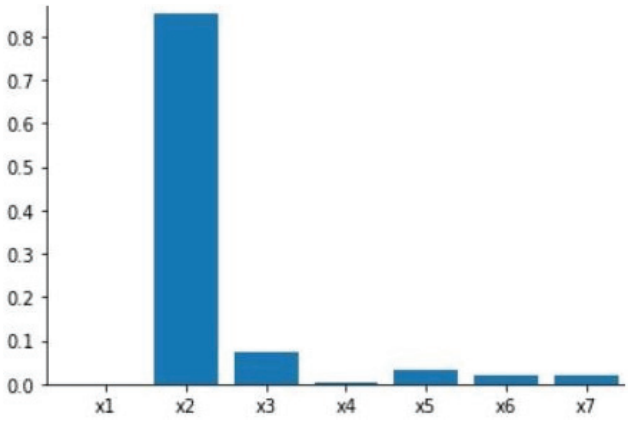


Fig. 3. Effect of Meteorological Factors on Solar Power Determined Using Random Forest Regression

Seven input variables are fed to the RF and their relative importance with respect to solar power is plotted in the form of bar-graph as shown in Fig. 3. Here features are denoted as xt (where t ∈ [1, 7]) where: x1: is_daylight; x2: dist_to_solar_noon; x3: cloud_cover; x4: visibility; x5: rel_humidity; x6: avg_ws_period and x7: avg_bp_period. Table 2 displays the list of all input features ranked from 1 to 7 based on their respective importance value distributed in the range of 0 to 1. All importances together sum up to 1. The top four ranked features, highlighted in table 2, arranged in descending order of their importance are selected as the best inputs. The relatively bigger bars representing the importance values (along the y-axis), are those of the four most effective parameters as shown in Fig. 3.

TABLE II. RANKING OF IMPORTANT PARAMETERS

Rank	Input Parameter	Importance Value
1	dist_to_solar_noon	0.8539
2	cloud_cover	0.0748
3	rel_humidity	0.0301
4	avg_bp_period	0.0198
5	avg_ws_period	0.0186
6	visibility	0.0025
7	is_daylight	0.000000514

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:59:05 UTC from IEEE Xplore. Restrictions apply.

E. Cross-Validation

The objective of this step is to enhance accuracy and performance of the LSTM model by applying hyperparameter tuning in order to secure the optimal values of various hyperparameters. In the cross-validation step, processed data is firstly partitioned sequentially into three sections, videlicet training, validation and test datasets. Here 80% of the dataset is utilized for training, 10% for validation and 10% for test. Batch size is acted upon in order to secure its best value. Impact of two values of window length (24 and 12) and three different optimizers - Adam, Adamax, RMSProp – is further scrutinized as part of the tuning process.

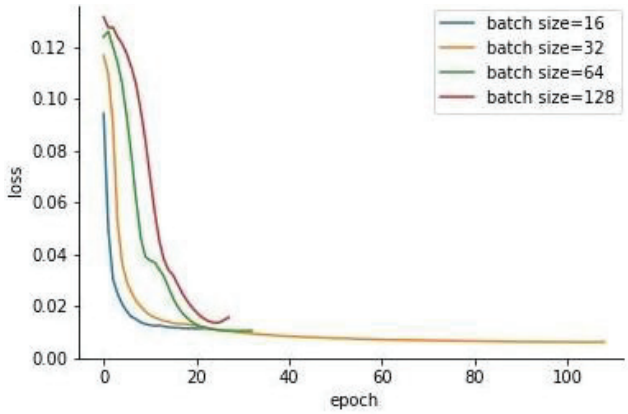


Fig. 4. Tracking of Validation Loss over Epochs for Different Values of Batch Size

Here, manual tuning was utilized to carry out hyperparameter tuning. For a fixed window length, performance of four different values of batch size (16, 32, 64 & 128), under the effects of aforementioned optimizers was evaluated. This process was repeated for another predetermined value of the window length yielding the optimal model which outperforms the other 24 models based on its low validation MAE. Performance of the selected model is then tested over an unknown set termed the test dataset. Fig. 4 illustrates the tuning results of batch size wherein a plot of loss evolution in the function of epochs is studied for different values of batch size until the model converges into results similar to the observed data. The best hyperparameter values are tabulated and highlighted in table 3.

TABLE III. SUMMARY OF HYPERPARAMETER TUNING

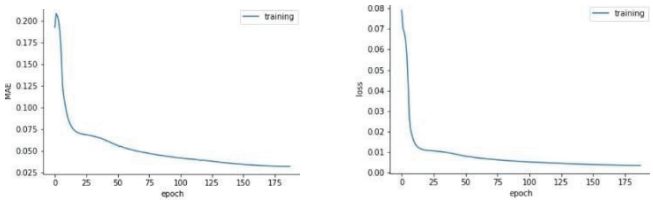
Model No.	Batch Size	Optimizer	Window Length	Validation MAE
1	16	Adam	12	0.0484
2	16	Adamax	12	0.0517
3	16	RMSProp	12	0.0547
4	32	Adam	12	0.0412
5	32	Adamax	12	0.0547
6	32	RMSProp	12	0.0532
7	64	Adam	12	0.0562
8	64	Adamax	12	0.0761
9	64	RMSProp	12	0.0503
10	128	Adam	12	0.0599
11	128	Adamax	12	0.0906
12	128	RMSProp	12	0.0540
13	16	Adam	24	0.0693
14	16	Adamax	24	0.0506
15	16	RMSProp	24	0.0617
16	32	Adam	24	0.0396
17	32	Adamax	24	0.0429
18	32	RMSProp	24	0.0440
19	64	Adam	24	0.0666
20	64	Adamax	24	0.0513
21	64	RMSProp	24	0.0542
22	128	Adam	24	0.0900
23	128	Adamax	24	0.0603
24	128	RMSProp	24	0.0525

F. Training the LSTM Network

Post the cross-validation step, the optimal hyperparameters as listed in table 4 are used to construct the LSTM networks on the training set. The training MAE and loss values were calculated to be 0.0306 and 0.0032 respectively. The training progress report showcases how MAE varies in the function of epochs as shown in Fig. 5(a) and Fig. 5(b), explains how the loss function has evolved over increasing epochs. It can be concluded that both the loss function and MAE converge when epochs are around 175. Designing of the LSTM networks incorporates a few aspects such as overfitting which is countered using the class callbacks of Keras library in Python. Furthermore, Adam optimizer has been utilized for compilation. Lastly, functioning of early stopping condition employing the validation set can be understood in the sense that it puts a halt to training once validation loss stops changing for a certain number of epochs. Eventually, forecasting is possible following the construction of LSTM networks.

TABLE IV. PARAMETERS OF LSTM MODEL

Parameters	Value
No. of LSTM layers	1
Optimizer	Adam
No. of training epochs	200
Learning Rate	0.0001



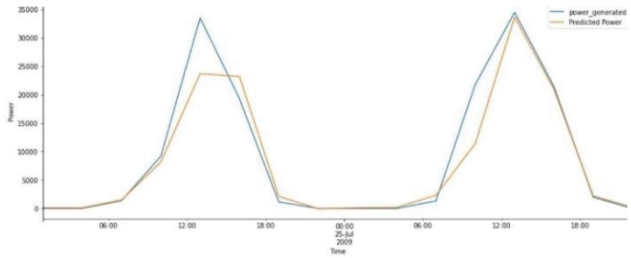
(a) MAE vs Epochs (b) Loss Function vs Epochs

Fig. 5. Training Progress Report and Evolution of Loss Function

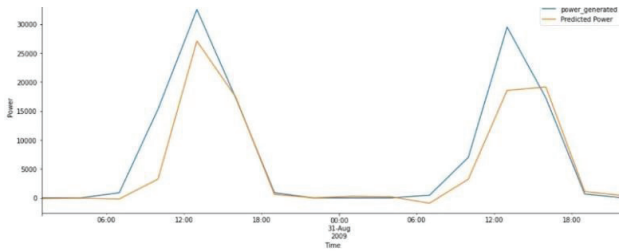
V. RESULTS AND DISCUSSION

This section presents forecasting results of the suggested method using performance metric mean absolute error (MAE). A feature engineering technique employed in this research selects the four most important features - distance to solar noon, cloud cover, relative humidity and average barometric pressure period. Here, solar power has been forecasted on a singlestep sequential basis for a three-hour prediction horizon reflecting the application of very short- term forecasting. In all the plots as shown in Fig. 6 to Fig. 8, actual power (here power_generated) is represented by a blue curve whereas an orange curve represents the predicted values of power. A graphical comparison between the actual generated power and predicted power delineates the precision of any forecasting method.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:59:05 UTC from IEEE Xplore. Restrictions apply.



(a) Forecast on Validation Set



(b) Forecast on Test Set

Fig. 6. Multivariate Forecasts after Feature Engineering

TABLE V. COMPARISON OF UNIVARIATE AND MULTIVARIATE APPROACH

Dimensionality	Validation MAE
Univariate	0.0498
Multivariate	0.0396

Mean absolute error in the entire forecasting period for the proposed multivariate approach (on test set) is found to be relatively small, MAE = 0.0565. Further, error evaluation is performed on the validation set to draw a comparative analysis between univariate and multivariate ways of forecasting as shown in table 5. Lesser value of the validation MAE for multivariate approach exhibit its supremacy over the univariate approach as demonstrated in Fig. 6(a) and Fig. 7 respectively.

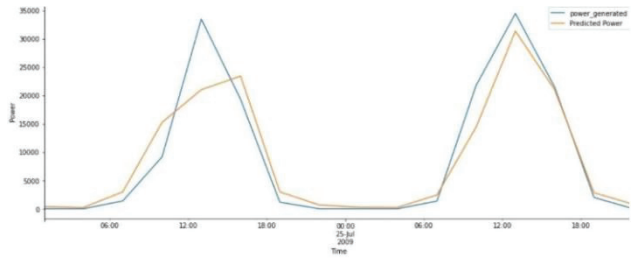


Fig. 7. Univariate Forecast (Validation Set)

Fig. 8 displays the case where selection of variables was not carried out with the MAE on the validation set being 0.0403 which is clearly greater than the error value after the application of selection technique (MAE = 0.0396), asserting the importance of feature engineering on the accuracy of forecasts. Fig. 6 shows that the actual and predicted values are closely related and evince the presented LSTM model's forecasting ability on a very short-term basis. As far as the accuracy is concerned, there is still room for improvement which is possible with the usage of other meteorological parameters such as irradiance, wind direction and ambient temperature.

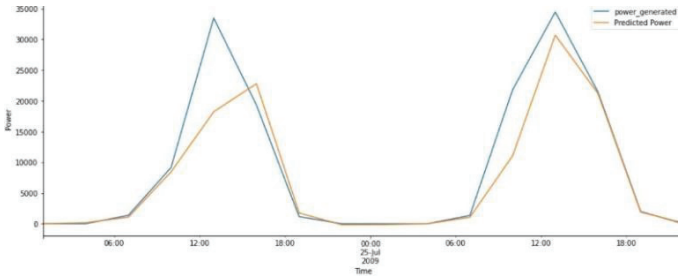


Fig. 8. Multivariate Forecast before Feature Engineering

VI. CONCLUSION

Variability induced by the intermittent solar energy source has an acute influence on the generated solar power quality with plants experiencing a drop in power generation. Accurate short-term power predictions are key for utility companies which use electricity price forecasts extensively to maximize their monetary benefits and create a schedule for generated electricity concerned with user demand. Using a chain-like structure, LSTM mitigates the effect of vanishing or exploding gradient problem associated with RNN (recurrent neural network). Being a promising and effective technique, this paper proposes a LSTM-based singlestep forecasting approach to solar power. Monthly data belonging a power generation system installed in Berkeley, California is utilized in this study. Feature engineering concludes distance to solar noon, cloud cover, relative humidity and average barometric pressure period as the important input parameters. The MAE for the multivariate approach came out to be 0.0565. Promising results advocate the use of LSTM for multivariate very short-term solar PV power forecasting and indicate its superior performance compared to univariate approach. However, non-inclusion of certain other meteorological variables like irradiance, wind direction and ambient temperature can be regarded as a limitation of this study. We plan on extending this work in future by using a more robust method like grid-search for the purpose of tuning over other parameters like epochs and learning rate. Individual parameters of optimizers will also be modified for a deeper investigation of tuning effects. Furthermore, future studies will take on the comparative evaluation of LSTM with other neural network architectures like GRU (gated recurrent unit). As a matter of future work, a multivariate ensemble framework involving extra trees (ET) and random forest (RF) as ML techniques will be studied on PV systems.

REFERENCES

[1] IEA (2020), Renewables 2020, IEA, Paris. [Online]. https://www.iea.org/reports/renewables-2020, accessed March 18, 2021.

[2] V. Shanthan Reddy and Dikshi Arora "Renewable Energy PPA's in India: A Kick at the Can or a Fly in the Ointment?". [Online]. https://ijpiel.com/index.php/2020/10/23/renewable- energy-ppas-in-india-a-kick-at-the-can-or-a-fly-in-the- ointment/, accessed March 18, 2021.

[3] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F.J. Martinez- de-Pison, F. Antonanzas-Torres, Review of photovoltaic power forecasting, Solar Energy, Volume 136, 2016, Pages 78-111, ISSN0038-092X,

[4] https://doi.org/10.1016/j.solener.2016.06.069.

[5] Salpakari, J.; Mikkola, J.; Lund, P.D. Improved flexibility with large-scale variable renewable power in cities through optimal demand side management and power-to-heat conversion. Energy Convers. Manag. 2016, 126, 649–661

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:59:05 UTC from IEEE Xplore. Restrictions apply.

[6] Kushagra Bhatia, Rajat Mittal, Jyothi Varanasi, M.M. Tripathi, An ensemble approach for electricity price forecasting in markets with renewable energy resources, Utilities Policy, Volume 70, 2021, 101185, ISSN 0957-1787, https://doi.org/10.1016/j.jup.2021.101185.

[7] Bacher, P, Madsen, H & Nielsen, HA 2009, 'Online short-term solar power forecasting', Solar Energy, vol. 83, no. 10, pp. 1772-1783. https://doi.org/10.1016/j.solener.2009.05.016.

[8] A. S. B. Mohd Shah, H. Yokoyama and N. Kakimoto, "High-Precision Forecasting Model of Solar Irradiance Based on Grid Point Value Data Analysis for an Efficient Photovoltaic System," in IEEE Transactions on Sustainable Energy, vol. 6,no. 2, pp. 474-481, April 2015, doi: 10.1109/TSTE.2014.2383398.

[9] Yang, C., Thatte, A., & Xie, L. (2015). Multitime-Scale Data- Driven Spatio-Temporal Forecast of Photovoltaic Generation. IEEE Transactions on Sustainable Energy, 6, 104- 112.

[10] James, G.; Witten, D.; Hastie, T.; Tibshirani, R. An Introduction to Statistical Learning; Springer: Berlin/Heidelberg, Germany, 2013; Volume 112.

[11] Gordon Reikard, Predicting solar radiation at high resolutions: A comparison of time series forecasts, Solar Energy, Volume 83, Issue 3, 2009, Pages 342-349, ISSN 0038-092X, https://doi.org/10.1016/j.solener.2008.08.007.

[12] Yanting Li, Yan Su, Lianjie Shu, An ARMAX model for forecasting the power output of a grid connected photovoltaic system, Renewable Energy, Volume 66, 2014, Pages 78-89, ISSN 0960-1481,

[13] https://doi.org/10.1016/j.renene.2013.11.067.

[14] Cyril Voyant, Gilles Notton, Soteris Kalogirou, Marie-Laure Nivet, Christophe Paoli, Fabrice Motte, Alexis Fouilloy, Machine learning methods for solar radiation forecasting: A review, Renewable Energy, Volume 105, 2017, Pages 569-582, ISSN 0960-1481,

[15] https://doi.org/10.1016/j.renene.2016.12.095.

[16] Amit Kumar Yadav, Hasmat Malik, S.S. Chandel, Selection of most relevant input parameters using WEKA for artificial neural network based solar radiation prediction models, Renewable and Sustainable Energy Reviews, Volume 31, 2014, Pages 509-519, ISSN 1364-0321, https://doi.org/10.1016/j.rser.2013.12.008.

[17] A. Gensler, J. Henze, B. Sick and N. Raabe, "Deep Learning for solar power forecasting — An approach using AutoEncoder and LSTM Neural Networks," 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Budapest, Hungary, 2016, pp. 002858-002865, doi: 10.1109/SMC.2016.7844673.

[18] Bouzerdoum, M., Mellit, A., & Pavan, A.M. (2013). A hybrid model (SARIMA-SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant. Solar Energy, 98, 226-235.

[19] Yang, H-T., Huang, C. M., Huang, Y. C., & Pai, Y. S. (2014). A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output. IEEE Transactions on Sustainable Energy, 5(3),917-926.

[20] [6802349]. https://doi.org/10.1109/TSTE.2014.2313600.

[21] Xiangyun Qing, Yugang Niu, Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM, Energy, Volume 148, 2018, Pages 461-468, ISSN 0360-5442, https://doi.org/10.1016/j.energy.2018.01.177.

[22] Ahmad Alzahrani, Pourya Shamsi, Cihan Dagli, Mehdi Ferdowsi, Solar Irradiance Forecasting Using Deep Neural Networks, Procedia Computer Science, Volume 114, 2017, Pages 304-313,ISSN1877-0509, https://doi.org/10.1016/j.procs.2017.09.045.

[23] Abdel-Nasser, M., Mahmoud, K. Accurate photovoltaic power forecasting models using deep LSTM-RNN. Neural Comput & Applic 31, 2727–2740 (2019). https://doi.org/10.1007/s00521- 017-3225-z

[24] Hall, M., & Smith, L.A. (1999). Feature Selection for Machine Learning: Comparing a Correlation-Based Filter Approach to the Wrapper. FLAIRS Conference.

[25] M. Q. Raza, N. Mithulananthan, J. Li, K. Y. Lee and H. B. Gooi, "An Ensemble Framework for Day-Ahead Forecast of PV Output Power in Smart Grids," in IEEE Transactions on Industrial Informatics, vol. 15, no. 8, pp. 4624-4634, Aug. 2019, doi: 10.1109/TII.2018.2882598.

[26] Pascanu, R., Mikolov, T. & Bengio, Y.. (2013). On the difficulty of training recurrent neural networks. Proceedings of the 30th International Conference on Machine Learning, in PMLR 28(3):1310-1318

[27] Graves A. (2012) Supervised Sequence Labelling. In: Supervised Sequence Labelling with Recurrent Neural Networks. Studies in Computational Intelligence, vol 385. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978- 3-642-24797-2_2

[28] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[29] Thomas Fischer, Christopher Krauss, Deep learning with long short-term memory networks for financial market predictions, European Journal of Operational Research, Volume 270, Issue 2, 2018, Pages 654-669, ISSN 0377-2217, https://doi.org/10.1016/j.ejor.2017.11.054.

[30] Kaggle Solar Power Generation. [Online]. Available: https://www.kaggle.com/vipulgote4/solar-power-generation, accessed March 1, 2021.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:59:05 UTC from IEEE Xplore. Restrictions apply.

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

Wireless Energy Harvesting Cognitive Radio Network under Finite-Capacity Battery for Nakagami-m Fading

Nayanika Biswas
Electronics and Communication Engineering Dept.
Delhi Technological University
New Delhi, India
nayanikabiswas_2k17ec109.dtu.ac.in

Piyush Tewari*
Electronics and Communication Engineering Dept.
Delhi Technological University
New Delhi, India
piyushtewari@dtu.ac.in

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498306

***Abstract*—We present a radio frequency(RF) based energy harvesting(EH) cognitive radio network(CRN) with finite-capacity battery at the secondary source(SS). All channels in the network undergo Nakagami-m fading. The secondary network's(SN) performance is analysed using outage probability and throughput. To obtain a simplified analytical model, we derive the expressions for the upper bound of the outage probability and the lower bound for the throughput of the SN. The analytical expressions for the PDF and CDF of harvested energy of the SN are also derived. The impact of finite battery constraint on the performance of the SN is investigated and analytical expressions for outage probability and throughput are derived. Results show that for an optimal battery capacity, the performance of the SN is identical to that of the infinite-capacity battery assisted CRN. The paper also presents a method to obtain the optimal battery capacity of the CRN under Nakagami-m fading. All analytical results are verified against Monte-Carlo simulations in MATLAB.**

***Index Terms*—Cognitive Radio, Wireless Energy Harvesting, Finite-Capacity Battery Model, Nakagami-m Fading Channel**

I. Introduction

With the rise of Internet of Things(IOT), there is an increase in the use of wireless devices used by individuals. A huge amount of spectrum is required to support all these wireless devices. Spectrum is a natural resource licensed by Governments. Businesses have opted for digitisation and shifted from brick and mortar setups. However, this natural resource is scarce and under utilized. Spectrum Holes are frequency bands in a spectrum over whose entirety the secondary user can transmit signal without causing interference at the primary receiver. Cognitive Radios are being introduced to tackle this shortage and under utilisation of spectrum by taking advantage of these spectrum holes [7]. Cognitive radios Networks(CRNs) allow opportunist spectrum sharing and spectrum sensing systems called Secondary Users(SUs)(also known as unlicensed users) that dynamically adapt to unused spectrums allocated to licensed users called Primary Users(PU) by adapting its parameters in real-time. [7]- [6]. The performance and design of CRNs for relay-assisted and non-relay assisted networks have been investigated in the scientific literature [1], [4].

Another major concern is that with the rise in wireless devices, there is an increase in their power requirement. To prolong the lifetime of energy-constrained wireless networks, make them self-sustainable, and promote green communication, energy harvesting techniques that harness energy from sources like heat, wind, light and radio-waves are being studied [14], [12]. Recently, EH due to ambient radio frequency(RF) signals has caught the attention of the research community as it provides energy self sufficiency conveniently to a low powered communication system [11], [8], [2]. With increase demand and technological advances in low power devices in academia and industry, wireless energy harvested[WEH] from RF signals will be practically realisable, specially for wireless sensor network(WSN) nodes. [8], [15].

Commonly used EH secondary receiving(SR) architecture are power splitting(PS) architecture and time splitting(TS) architecture [14]. In PS, the received signal power is split into two fractions, one for EH and the other for information retrieval [14], [8]. Whereas, in TS the energy harvesting-information transmitting(EH-IT) cycle is split into two fractions, one for EH and information retrieval by the receiver node and other for IT [8], [15].

In [13], an underlay non relay assisted secondary network(SN) is analyzed using outage probability and the results show that the energy harvested by the SN from the PU transmitter's RF signals is sufficient for information transmission of SN as well as relaying information for the PN. In [5], the outage probability and throughput of an underlay relay assisted CRN with TS protocol is analysed. [5] uses Rayleigh fading channel model and multiple PU transmitters and receivers.In [10], an non-relay underlay EH CRN is proposed and the impact of multiple energy- constrained receivers in SN is studied. Most works on CRNs that provide a good understanding of the subject assume channels to undergo Rayleigh fading [13], [5]. They also assume infinite storage capacity at the energy harvesting nodes of the SN [13], [5]. While Rayleigh fading channels simplify analytical solutions, but they represent limited practical scenarios [3]. Similarly, the rechargeable battery at the WEH node of the SN has a finite battery capacity. Nakagami-m fading channels represent a variety of scenarios whose severity can be tuned using

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:02:09 UTC from IEEE Xplore. Restrictions apply.

the shape parameter. Further, for m = 1, m $<$ 1, and m $>$ 1, a Nakagami-m fading channel can be modelled as a Rayleigh fading channel, Hoyt fading channel, and Rice fading channel respectively [3]. In [3], the outage probability of dual-hop cognitive amplify-and-forward (AF) relay-assisted CRN undergoing Nakagami-m fading is studied. It assumes a finite maximum transmit power at SN, provides comparison between two proposed transmit power constraints and shows that the irrespective of the transmit power constraint, the diversity order is strictly defined by the minimum fading severity between the two transmitting nodes at the SN. In [9], an overlay CRN is considered which nakagami-m fading channels, in which a secondary transmitter (ST) relays the information of a PU transmitter toward a PU receiver.

A. Motivation and Contribution

We noted a lack of analytical models in the technical literature for WEH-CRN under HU architecture, which studies the optimal battery capacity required at the SN, under Nakagami-m fading condition, for optimal performance. In order to study the impact of finite battery constraint on the performance of WEH-CRN along with other system parameters, the following contributions are made by the proposed work:

1) We derive analytical expressions for the upper bound of the outage probability as well as the lower bound for the throughput for the SN for a non-relay-assisted WEH underlay CRN with finite-capacity battery at SS and investigate its impact on the performance of SN under decode and forward protocol for delay-sensitive networks. The following constraints are enforced: 1) The harvested energy is limited by the maximum storage capacity of the rechargeable battery at SS; 2) The maximum transmit power at SS is limited by the energy harvested; 3) The maximum transmit power at SS is limited by the peak interference power threshold set by PU; 4) The interference power from multiple PU transmitters at SD limits reliable communication between SS and SD.
2) Novel expressions for the PDF and CDF for the harvested energy at the EH node with finite-capacity storage(i.e., SS) are derived.
3) The outage probability and throughput for SN are evaluated for different battery storage capacity values, and the optimum battery capacity is identified.
4) We show that the impact of residual harvested energy remaining in the battery on the performance of the SN is minimal.

II. Network Model

We consider a non-relay underlay CRN where an energy constrained SS uses the PU's licensed spectrum to transmit to the SD. The PN has N primary transmitters and M primary receivers. All PU transmitters are assumed to be clustered about a point due to their proximity. Similarly, all PU receivers too are clustered about a central point as they are closely located [5]. The SS only uses radio frequency signals from the PU transmitters to harvest energy. The harvested energy is stored in a finite capacity battery which can be modelled as a rechargeable super-capacitor with a switching circuit to alternate between the energy harvesting and signal transmission cycles of the CRN. As shown in figure 1, the coefficients $g_{S,i}$ for $i = 1, 2, ..., M$, represent the channel gains from the ith PU_{RX} to SS respectively. The coefficient h denotes the channel gain coefficient from SS to SD. Similarly, the coefficients $f_{S,j}$, $f_{D,j}$ for $j = 1, 2, , .., N$, represent the channel gains from the jth PU_{TX} to SS and SD respectively. All channels are modelled assuming the distance between the jth PU_{TX} and SS, and SD as $d_{S,j}$, $d_{D,j}$, respectively. The distance between ith PUrx and SS is denoted as $d_{S,i}$ respectively. And the distance between SS and SD are denoted as d4, respectively. The link gain realizations $|h|^2, |g_S|^2, |f_{S,j}|^2, |f_{D,j}|^2$ are gamma distributed with shape parameters $m_h, m_{g_{S,i}}, m_{f_{S,j}}, m_{f_{D,j}}$, and $\beta_h, \beta_{g_{S,i}}$, $\beta_{f_{S,j}}$, and $\beta_{f_{D,j}}$ respectively. $\beta_{f_{S,j}} = {d_{1,j}}^{-n}$, n being the path loss factor. $f_{S,j}$ $\forall j$ that link the PU transmitters and the SS are assumed to be identically distributed. Similarly, $f_{D,j}$ $\forall j$ that link PU transmitters and SD follow identical distribution. Moreover, identical distribution is assumed for the link gain from SS to PU receivers due to close proximity of all the transmitters and all the receivers about their respective central points. However, $f_{S,j}$, $f_{D,j}$, and $g_{S,i}$ are non-identically distributed wrt each other. The energy harvesting and information transmission is

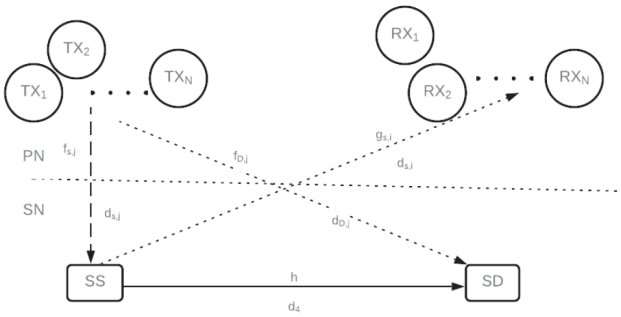


Fig. 1. System Model for WEH CRN. The SS has finite-capacity battery.

performed in alternate cycles, following time splitting architecture. T is the total time period for one cycle and α is the ratio at which every cycle is split into two sub cycles, one for harvesting energy from the PU transmitters and other for information transmission(IT).

The energy harvested at the SS under infinite battery capacity is

$$E_{hsi} = \eta\alpha \mathrm{P}_{\mathrm{PU_{tx}}} \sum_{j=1}^{N} |f_{1,j}|^2 \mathrm{T} \tag{1}$$

where η represents the energy conversion coefficient, $P_{PU_{tx}}$ denotes transmitted power from all PU transmitters and $|f_{1,j}|$ is Nakagami-m distributed with the PDF

$$f(x) = \frac{2m^m x^{2m-1} e^{\frac{-mx^2}{\beta_{1,j}}}}{\Gamma(m)\beta_{1,j}^m} \tag{2}$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:02:09 UTC from IEEE Xplore. Restrictions apply.

where $\beta_{1,j} = E[X^2]$ and $\Omega = \frac{\beta_{1,j}}{m}$ The link gain realisation $|f_{1,j}|^2$ follows gamma distribution whose PDF and CDF are defined as

$$f(x) = \frac{x^{m-1} e^{\frac{-x}{\Omega}}}{\Omega^m \Gamma(m)} \tag{3}$$

$$F(x) = \frac{\gamma(m, \frac{x}{\Omega})}{\Gamma(m)} \tag{4}$$

where $\gamma(m,x) = \int_0^x t^{x-1} e^{-t} dt$ is the lower incomplete gamma function.

Due to the finite-battery constraint the energy harvested at the SS is

$$E_{hs} = \min(E_{hsi}, B_{max}) \tag{5}$$

where B_{max} is the maximum battery capacity energy at SS.

The transmit power at SS is

$$P_s = \min\left(\frac{E_{hs}}{(1-\alpha)T}, \frac{P_I}{\min\{|g_{1,i}|^2\}}\right) \tag{6}$$

where P_I is the maximum permissible interference power at PU_{rx} due to SS. In a transmission cycle, when the term $\frac{E_{hs}}{(1-\alpha)T}$ is larger than the term $\frac{P_I}{\min\{|g_{1,i}|^2\}}$, the maximum transmitted energy is limited by the latter term thus giving rise to residual energy. This residual energy remaining in the battery is considered to be negligible for obtaining the upper bound on the outage probability of the SN. To avoid interference at PU receivers due to high power transmitted by the SS, the peak interference constraint is imposed such that it does not exceed P_I.

The signal to interference ratio (SIR) at SS is defined as

$$\Gamma_S = \min\left(\frac{E_{hs}}{(1-\alpha)T}, \frac{P_I}{Y}\right)\frac{X}{Z_2} \tag{7}$$

$Z_1 = P_{PU_{tx}} \sum_{j=1}^{N} |f_{S,j}|^2, Z_2 = P_{PU_{tx}} \sum_{j=1}^{N} |f_{D,j}|^2, X = |h|^2, Y = \min\{|g_{S,i}|^2\}$

We evaluate the performance of the SN using two system performance metrics, outage probability and network throughput.

Outage Probability(OP) is defined as the probability that the SIR at SS is below a threshold value γ_{th} [5]. We consider the presented non-relay assisted underlay network to be in outage iff outage is suffered by the SS-SD link. Its expression is given by

$$P_{out}(\gamma_{th}) = 1 - Pr\{\Gamma_S \leq \gamma_{th}\} \tag{8}$$

where Γ_S is the SIR at the SD. Throughput of a network is stated as the amount of data received at the receiver end after transmission through the communication link[9]. We evaluate the throughput in delay-sensitive mode. In this mode, the SS transmits information at a fixed rate R_{ds} to SD. The throughput in this mode is given as[14]

$$\tau_{ds} = (1-\alpha) R_{ds} (1 - P_{out}(\gamma_{th})) \tag{9}$$

where τ_{ds} is the rate of successful transmissions during the transmission cycle, given that the SIR is more than the threshold γ_{th} at SD.

III. Performance Analysis

We evaluate the performance of the network based on certain thresholds above which the network is considered reliable. In this section, we derive the analytical solution for the upper bound for OP of the proposed CRN model.

The OP can also be written as

$$P_{out}(\gamma_{th}) = \mathrm{I}_1 + \mathrm{I}_2 - \mathrm{I}_3 \tag{10}$$

where $I_1 = P(V < \zeta), I_2 = P(\frac{P_I}{Y} < \zeta), I_3 = P(V < \zeta, \frac{P_I}{Y} < \zeta), \zeta = \frac{\gamma_{th} Z_2}{X}$ and $V = \frac{E_{hs}}{(1-\alpha)T}$

The PDF and CDF of the E_{hs} is required to derive the mathematical expression of the upper bound for outage probability using eq. (8). We obtain the CDF and PDF of E_{hsi} as,

$$F_{E_{hsi}} = \frac{\Gamma\{\sum_{j=1}^{N} m_{S,j}, \frac{x}{c\Omega}\}}{\Gamma\{\sum_{j=1}^{N} m_{S,j}\}} \tag{11}$$

$$f_{E_{hsi}} = gamma\left(\sum_{j=1}^{N} m_{S,j}, c\Omega\right) \tag{12}$$

where $\Gamma(.,x)$ is the upper incomplete gamma function, and $gamma(.,.)$ denotes the gamma distribution and $c = \eta P_{PU_{TX}} T$. As seen in eq (5), the energy stored in the battery is equal to E_{hsi} when it is less than B_{max}, else it is equal to B_{max}, as B_{max} acts as the ceiling for the maximum harvested energy that can be stored in the battery. Thus the CDF can be written as

$$Pr(E_{hs} < ehs | ehs < B_{max}) = Pr(E_{hsi} < ehs)$$

and

$$Pr(E_{hs} < ehs | ehs > B_{max} = 1 \mathrm{as} Pr(E_{hs} < B_{max}) = 1 \tag{13}$$

The PDF can be derived by piece-wise differentiation of the CDF. Using eq (13), the CDF of E_{hs} can be expressed as

$$F_{E_{hs}}(ehs) = \frac{\Gamma\{\sum_{j=1}^{N} m_{S,j}, \frac{x}{c\Omega}\}}{\Gamma\{\sum_{j=1}^{N} m_{S,j}\}} [u(ehs) - u(ehs - B_{max})] + u(ehs - B_{max}) \tag{14}$$

Thus the optimal B_{max} value required to achieve performance of SN that is identical to that of infinite-capacity battery assisted CRN can be obtained using the following relation

$$F_{E_{hs}}(B_{max}) = \frac{\Gamma\{\sum_{j=1}^{N} m_{S,j}, \frac{B_{max}}{c\Omega}\}}{\Gamma\{\sum_{j=1}^{N} m_{S,j}\}} = 0.99 \tag{15}$$

$$B_{max} = \gamma^{-1}\left(\sum_{j=1}^{N} m_{S,j}, 0.99\Gamma\{\sum_{j=1}^{N} m_{S,j}\}\right)(c\Omega) \tag{16}$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:02:09 UTC from IEEE Xplore. Restrictions apply.

where $\gamma^{-1}(.,.)$ denotes the inverse of lower incomplete gamma function. Using definitions given by eq. (10), we solve for I_1,

$$\begin{aligned}
I_1 &= Pr\{min(B_{max}, kZ_1) < \zeta(1-\alpha)T\} \\
&= Pr\{X < \frac{\gamma_{th} Z_2 (1-\alpha)T}{B_{max}}\} \times Pr\{Z_1 > \frac{B_{max}}{k}\} \\
&+ Pr\{X < \frac{\gamma_{th} Z_2 (1-\alpha)T}{kZ_1} \cap Z_1 < \frac{B_{max}}{k}\} \\
&= \left[\frac{1}{C_1} \times \left(\frac{C_2^{m_h}\Gamma(m_h + M_{jD})}{m_h (C_2 + \frac{1}{\Omega'_{jD}})^{M_{jD}+m_h}}\right)\right. \\
&\left. \times F_{21}\left(1, m_h + M_{jD}; m_h + 1; \frac{C_2}{C_2 + \frac{1}{O'_{jD}}}\right)\right] \\
&\times \left(1 - \frac{\gamma\left(M_{jS}, \frac{B_{max}}{k\Omega'_{jS}}\right)}{\Gamma(M_{jS})}\right) \\
&+ \int_0^{\frac{B_{max}}{k}} \left[\frac{1}{C_4}\left(\frac{C_3^{m_h}\Gamma(m_h + M_{jD})}{m_h (C_3 + \frac{1}{\Omega'_{jD}})^{M_{jD}+m_h}}\right)\right. \\
&\times F_{21}\left(1, m_h + M_{jD}; m_h + 1; \frac{C_3}{C_3 + \frac{1}{O'_{jD}}}\right) \\
&\left. \times e^{\frac{-z_1}{\Omega'_{jS}}} z_1^{M_{jS}-1} dz_1\right]
\end{aligned} \tag{17}$$

where $k = \eta\alpha T$, $C_1 = \Gamma(m_h)\Gamma(M_{jD})\left(\Omega'_{jD}\right)^{M_{jD}}$, $C_2 = \frac{\gamma_{th}(1-\alpha)T}{B_{max}\Omega_h}, C_3 = \frac{\gamma_{th}(1-\alpha)T}{Z_1 k \Omega_h}$, $C_4 = C_1\Gamma(M_{jS})\left(\Omega'_{jS}\right)^{M_{jS}}$,

and $F_{21}\left(a, b; c; z\right)$ is a hypergeometric function.

Solving for I2,

$$\begin{aligned}
I_2 &= Pr\{\frac{P_I}{Y} < \zeta\} \\
&= Pr\{X < \frac{\gamma_{th} Z_2 Y}{P_I}\} \\
&= F_{Z_2 Y}\left[\frac{\gamma\left(m_h, \frac{\gamma_{th} Z_2 Y}{P_I \Omega_h}\right)}{\Gamma(m_h)}\right] \\
&= \int_0^{\infty} \left[\frac{1}{C_6} \times \left(\frac{C_5^{m_h}\Gamma(m_h + M_{jD})}{m_h (C_5 + \frac{1}{\Omega'_{jD}})^{M_{jD}+m_h}}\right)\right. \\
&\times F_{21}\left(1, m_h + M_{jD}; m_h + 1; \frac{C_5}{C_5 + \frac{1}{O'_{jD}}}\right) \\
&\left. \times \frac{\gamma\left(m_{iS}, \frac{y}{\Omega iS}\right)^{M-1}}{\Gamma(m_{iS})^{M-1}} \times e^{\frac{-y}{\Omega'_{iS}}} y^{m_{iS}-1} dy\right]
\end{aligned} \tag{18}$$

where $C_5 = \frac{\gamma_{th} Y}{P_I \Omega_h}, C_6 = \frac{C_1 \Gamma(m_{iS})(\Omega_{iS})^{M_{iS}}}{M}$

Solving for I_3,

$$\begin{aligned}
I_3 &= P(V < \zeta, \frac{P_I}{Y} < \zeta), \\
&= \int_{z_2=0}^{\infty} \int_{x=0}^{\infty} I_{31} I_{32} f_{Z_2}(z_2) f_X(x) dz_2 d_x
\end{aligned} \tag{19}$$

as I_{31} and I_{32} are independent of each other wrt x and z_2. From eq (13)-(14), I_{31} can be simplified as

$$\begin{aligned}
I_{31} &= Pr\{E_{hsi} < ehs | ehs < B_{max}\} \\
&+ Pr\{E_{hsi} < ehs | ehs > B_{max}\} \\
&\text{`} = F_{E_{hsi}}(ehs)\left[u(ehs) - u(ehs - B_{max})\right] \\
&+ 1 \times u(ehs - B_{max}) \\
&= \frac{\gamma\left(M_{jS}, \frac{\gamma_{th} Z_2 (1-\alpha T}{X C_7 \Omega'_{jS}}\right)}{\Gamma(M_{jS})}\left[u(ehs) - u(ehs - B_{max})\right] \\
&+ 1 \times u(ehs - B_{max})
\end{aligned} \tag{20}$$

$ehs = \frac{\gamma_{th} Z_2 (1-\alpha)T}{X}, C_7 = \frac{\eta\alpha}{(1-\alpha)}$,

$$I_{32} = Pr\{Y > \frac{P_I X}{\gamma_{th} Z_2}\} = 1 - \left[\frac{\gamma\left(m_{iS}, \frac{P_I X}{\gamma_{th} Z_2 \Omega_{iS}}\right)}{\Gamma(m_{iS})}\right]^M \tag{21}$$

Using eq. (21) and (20) in (19), we write I_3 for $Z_2 < \frac{B_m axx}{\gamma_{th}(1-\alpha T}$

$$\begin{aligned}
I_3 &= \int_{x=0}^{\infty} \int_{z_2=0}^{\frac{x B_m ax}{\gamma_{th}(1-\alpha T}} \left(1 - \left[\frac{\gamma\left(m_{iS}, \frac{P_I X}{\gamma_{th} Z_2 \Omega_{iS}}\right)}{\Gamma(m_{iS})}\right]^M\right) \\
&\times \frac{\gamma\left(M_{jS}, \frac{\gamma_{th} Z_2 (1-\alpha T}{X C_7 \Omega'_{jS}}\right)}{\Gamma(M_{jS})} \\
&\times \frac{z_2^{M_{jD}-1} e^{\frac{-z_2}{\Omega'_{jD}}}}{\Gamma(M_{jD})(\Omega_{jD})'^{M_{jD}}} \frac{x^{m_h-1} e^{\frac{-x}{\Omega_h}}}{\Gamma(m_h)(\Omega_h)^{m_h}} \times dz_2 dx
\end{aligned} \tag{22}$$

and for $z_2 > \frac{B_{max} x}{\gamma_{th}(1-\alpha T}$

$$\begin{aligned}
I_3 &= \int_{x=0}^{\infty} \int_{z_2=0}^{\frac{x B_m ax}{\gamma_{th}(1-\alpha T}} \left(1 - \left[\frac{\gamma\left(m_{iS}, \frac{P_I X}{\gamma_{th} Z_2 \Omega_{iS}}\right)}{\Gamma(m_{iS})}\right]^M\right) \\
&\times \frac{z_2^{M_{jD}-1} e^{\frac{-z_2}{\Omega'_{jD}}}}{\Gamma(M_{jD})(\Omega_{jD})'^{M_{jD}}} \frac{x^{m_h-1} e^{\frac{-x}{\Omega_h}}}{\Gamma(m_h)(\Omega_h)^{m_h}} \times dx dz_2
\end{aligned} \tag{23}$$

IV. Results

In this section we analyse the performance of the SN of the proposed WEH-CRN model and discuss the implications of finite-capacity battery and other system parameters on it. The results of the analytical model for outage probability and throughput for the SN are presented and verified with Monte-Carlo simulations. Excellent agreement between the results of the simulations and analytical expressions is obtained.

The PU transmitters are clustered about (0,1) without loss of generality and the PU receivers are clustered around (1,1). SS and SD are positioned at (0,0) and (1,0) respectively. The

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:02:09 UTC from IEEE Xplore. Restrictions apply.

system parameters, until unless specifically mentioned, are assumed as $N = M = 3$, $\alpha = 0.5, \eta = 0.8$, and $T = 1$ [5]. The shape parameters, until unless specifically mentioned, are assumed to be $m_h = 2, m_{g_{S,i}} = 2, m_{f_{S,j}} = 3, m_{f_{D,j}} = 3$.

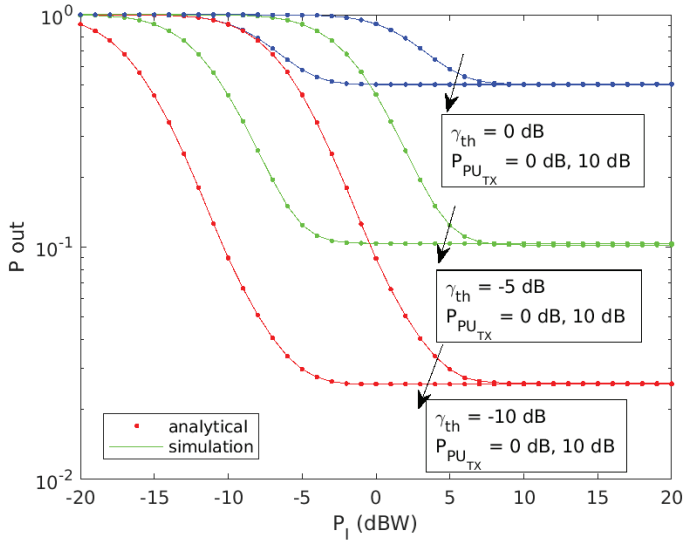


Fig. 2. OP(P_{out}) as a function of P_I for different $P_{PU_{TX}}$ and γ_{th}

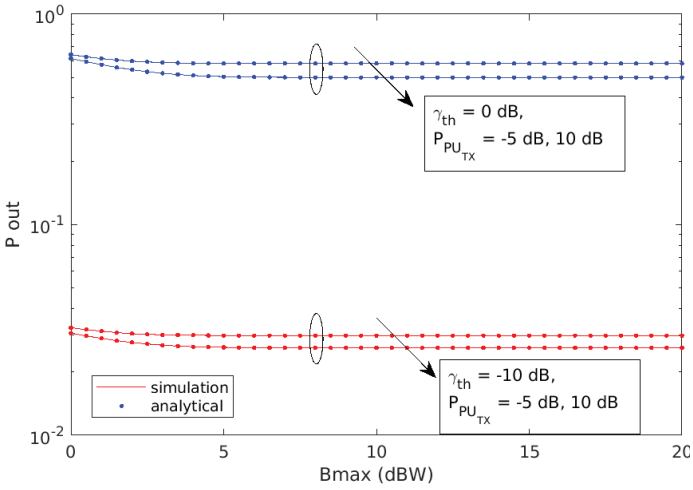


Fig. 3. OP(P_{out}) as a function of B_{max} for $P_{PU_{TX}} = 0dB$ and different γ_{th} and P_I

Fig 2 and 3 present the outage probability as a function of P_I and B_{max} respectively. We can see that the the OP decreases and ultimately saturates to a minimum as the P_I increases. This is because, as the permissible interference power increases beyond a value, the OP majorly depends on the harvested energy at the SS and becomes independent of P_I. Further, it can be observed that as the SIR threshold γ_{th} decreases, the OP decreases because of the relaxed constraint at SD for the successful detection of the transmitted signal. We can also see from the Fig 2 that for a larger $P_{PU_{tx}}$, the OP saturated for a higher P_I. This can be explained as the transmission power $P_{PU_{tx}}$ increases, the energy harvested at the SS increases, provided that the B_{max} is large enough, thus improving the SIR at the SD. We can see from Fig 3 that as the battery capacity increases the OP decreases and saturates to a minimum value where the OP becomes independent of the battery capacity. For a given γ_{th}, P_I and $P_{PU_{tx}}$, the system performance improves as B_{max} increases because it relaxes the constraint on maximum storage capacity of the finite battery at SS which in turn improves the SIR at the SD. It is also evident from Fig 3 that the system performance saturates to its minimum value after the optimal value of B_{max} at which it behaves identical to the asymptotic case for which the battery capacity is assumed to be infinite. The performance of SN after this optimal B_{max} value is independent of the battery capacity and depends on other system parameters like α and $P_{PU_{tx}}$. As seen in fig 3, the optimal B_{max} is observed to be 5.011 units. This agrees perfectly with the optimal B_{max} obtained from the eq (16), thus proving the validity of the eq (14).

Fig 4 illustrates the outage probability as a function of $P_{PU_{tx}}$. Here the OP increases with increase in $P_{PU_{tx}}$ for a fixed P_I until it reaches to unity and saturates. This is because increase in E_{hs} due to large $P_{PU_{tx}}$ is overpowered by the increase in interference at the SD. As the SIR is limited by P_I constraint, the OP eventually becomes independent of $P_{PU_{tx}}$.The result also shows that for a given γ_{th}, as the interference constraint relaxes, the OP saturates for a higher value of $P_{PU_{tx}}$. This is because increased tolerance to interference at the PU receivers allows storage of more harvested energy, given that B_{max} is large enough. The fig 4 also confirms the behaviour of OP wrt γ_{th}, i.e. OP increases with the increase in γ_{th}.

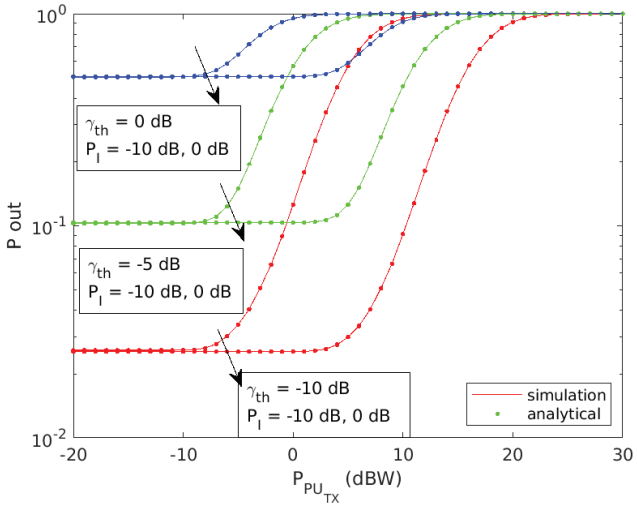


Fig. 4. OP(P_{out}) as a function of $P_{PU_{TX}} = 0dB$ for different γ_{th} and P_I

Fig 5 shows the delay-sensitive throughput of the SN as a function of P_I. For a fixed $P_{PU_{tx}}$, throughput increases and saturates at a maximum value as P_I increases. This is because

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:02:09 UTC from IEEE Xplore. Restrictions apply.

as the tolerance to interference at the PU receivers increases, the SS can transmit at higher power during the transmission cycle. For a higher value of $P_{PU_{tx}}$, the throughput increases and reaches its maximum value for a larger value of P_I. This is evident from the right shift in the throughput curve for $P_{PU_{tx}} = 10dB$.

Fig 6 shows that the OP for the proposed model is in close agreement with the OP curve which considers the effect of residual energy. This result proves that the proposed HU architecture based model provides a tight upper bound for the OP of the SN. Thus for a mathematically tractable solution, we can neglect the impact of residual energy.

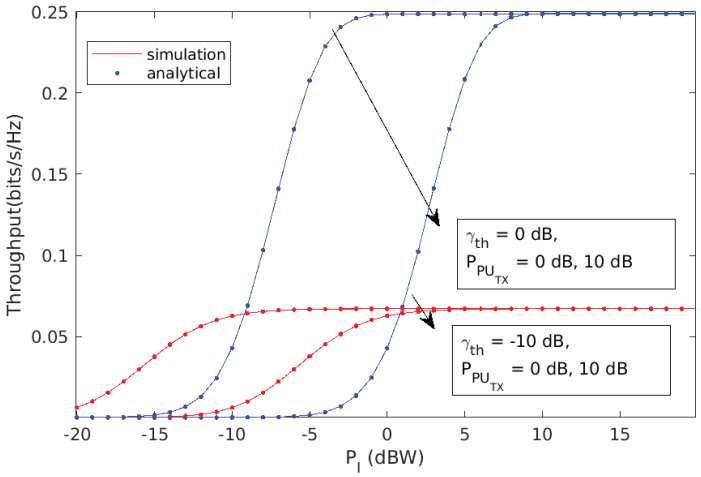


Fig. 5. Throughput as a function of P_I for different $P_{PU_{TX}}$ and γ_{th}

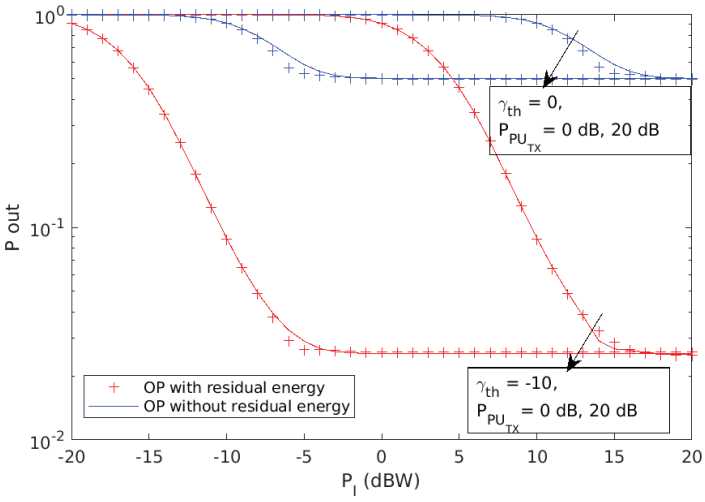


Fig. 6. Comparison Between the OP for WEH CRN considering residual energy in the battery with the proposed model

V. Conclusion

In this work, we presented the performance analysis for a WEH underlay non-relay CRN under HU protocol for Nakagami-m fading channels in terms of outage probability and throughput. The analytical expression for the upper bound of outage probability was obtained. We proposed novel PDF and CDF expressions for harvested energy at the SS to derive the expressions for outage probability and throughput. Our results show that optimal performance for the SN can be obtained with a finite-capacity battery, and in this region, the bound is in excellent agreement with the model considering the effects of residual energy remaining in the battery. We also presented a method to identify the optimum battery size corresponding to the optimal performance at the SN. Close agreement between the analytical and simulation results proves the validity of the proposed model.

References

[1] Mohamed M Abdallah, Ahmed H Salem, Mohamed-Slim Alouini, and Khalid A Qaraqe. Adaptive discrete rate and power transmission for spectrum sharing systems. *IEEE transactions on wireless communications*, 11(4):1283–1289, 2012.

[2] Hayder Al-Hraishawi and Gayan Amarasuriya Aruma Baduge. Wireless energy harvesting in cognitive massive mimo systems with underlay spectrum sharing. *IEEE Wireless Communications Letters*, 6(1):134–137, 2016.

[3] T. Q. Duong, D. B. da Costa, M. Elkashlan, and V. N. Q. Bao. Cognitive amplify-and-forward relay networks over nakagami- m fading. *IEEE Transactions on Vehicular Technology*, 61(5):2368–2374, 2012.

[4] YuanYuan He and Subhrakanti Dey. Throughput maximization in cognitive radio under peak interference constraints with limited feedback. *IEEE transactions on vehicular technology*, 61(3):1287–1305, 2012.

[5] Y. Liu, S. A. Mousavifar, Y. Deng, C. Leung, and M. Elkashlan. Wireless energy harvesting in a cognitive relay network. *IEEE Transactions on Wireless Communications*, 15(4):2498–2508, 2016.

[6] Yi Liu, Yan Zhang, Rong Yu, and Shengli Xie. Integrated energy and spectrum harvesting for 5g wireless communications. *IEEE Network*, 29(3):75–81, 2015.

[7] J. Mitola and G. Q. Maguire. Cognitive radio: making software radios more personal. *IEEE Personal Communications*, 6(4):13–18, 1999.

[8] Ali A Nasir, Xiangyun Zhou, Salman Durrani, and Rodney A Kennedy. Relaying protocols for wireless energy harvesting and information processing. *IEEE Transactions on Wireless Communications*, 12(7):3622–3636, 2013.

[9] B. V. Nguyen, H. Jung, D. Har, and K. Kim. Performance analysis of a cognitive radio network with an energy harvesting secondary transmitter under nakagami- m fading. *IEEE Access*, 6:4135–4144, 2018.

[10] L. Sibomana, H. Zepernick, and H. Tran. Wireless information and power transfer in an underlay cognitive radio network. In *2014 8th International Conference on Signal Processing and Communication Systems (ICSPCS)*, pages 1–7, 2014.

[11] Arooj Mubashara Siddiqui, Leila Musavian, Sonia Aissa, and Qiang Ni. Performance analysis of relaying systems with fixed and energy harvesting batteries. *IEEE Transactions on Communications*, 66(4):1386–1398, 2017.

[12] Kaya Tutuncuoglu and Aylin Yener. Optimum transmission policies for battery limited energy harvesting nodes. *IEEE Transactions on Wireless Communications*, 11(3):1180–1189, 2012.

[13] Zihao Wang, Zhiyong Chen, Ling Luo, Zixia Hu, Bin Xia, and Hui Liu. Outage analysis of cognitive relay networks with energy harvesting and information transfer. In *2014 IEEE international conference on communications (ICC)*, pages 4348–4353. IEEE, 2014.

[14] Fangchao Yuan, Shi Jin, Kai-Kit Wong, QT Zhang, and Hongbo Zhu. Optimal harvest-use-store design for delay-constrained energy harvesting wireless communications. *Journal of Communications and Networks*, 18(6):902–912, 2016.

[15] Rui Zhang and Chin Keong Ho. Mimo broadcasting for simultaneous wireless information and power transfer. *IEEE Transactions on Wireless Communications*, 12(5):1989–2001, 2013.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 07:02:09 UTC from IEEE Xplore. Restrictions apply.

Data and Text Mining

wiseR: An end-to-end structure learning and deployment framework for causal graphical models

Shubham Maheshwari[1], Khushbu Pahwa[2], Tavpritesh Sethi[1*]

[1] Indraprastha Institute of Information Technology, Okhla Industrial Estate, Phase III, Near Govind Puri Metro Station, New Delhi, Delhi 110020, India. [2] Delhi Technological University, Bawana Rd, Shahbad Daulatpur Village, Rohini, Delhi, 110042

*To whom correspondence should be addressed.

Abstract

Summary: Structure learning offers an expressive, versatile and explainable approach to causal and mechanistic modeling of complex biological data. We present *wiseR*, an open source application for learning, evaluating and deploying robust causal graphical models using graph neural networks and Bayesian networks. We demonstrate the utility of this application through application on for biomarker discovery in a COVID-19 clinical dataset.

Availability and implementation: Web application and source code of the R package are available at wiseR. Result dashboard at COVID19_Biomarkers

Contact: tavpriteshsethi@iiitd.ac.in | shubham14101@iiitd.ac.in

Supplementary information: Supplementary material is available for download online

1 Introduction

Discovering actionable knowledge from data is a major challenge in biology. Modern machine learning (ML) and deep learning have achieved major breakthroughs in fitting datasets. However, correlation does not imply causation, thus limiting the real world potential of these approaches to guide actions. Structure learning is an expressive approach for causal discovery from data (Drton and Maathuis 2017) and the visual nature of graphical structures makes them excellent for interactive decision making (Prosperi et al. 2020; Kleinberg and Hripcsak 2011; Richens, Lee, and Johri 2020). In small datasets with few variables, causal graphs can be built by using expert knowledge. However, this approach is not applicable in most real-world biological datasets which are high-dimensional. Fitting a unified graphical model without relying upon pairwise relationships is known to be an NP-hard problem. Recent advances in score-based and constraint-based approaches to learn Bayesian Networks (Scutari 2017) and more recently Graph ML and Graph Neural Networks (GNNs) (Yu et al. 2019) offer an opportunity to learn these directly from data. These approaches lead to a Directed Acyclic Graph (DAGs) with edge directions encoding causality or the flow of probabilistic influence, hence allowing causal discovery. The learned DAGs also provide a probabilistic model-based representation of knowledge which is appropriate for building actionable models from biomedical datasets. However, there are currently no open-source tools that allow end-to-end building, validating, learning actions and deploying these models.

In this paper, we present w*iseR*, an interactive R package integrating state-of-the-art structure learning methods including Graph Neural Networks, and Bayesian Networks for causal structure learning, visualization, probabilistic inference and model deployment as dashboards.

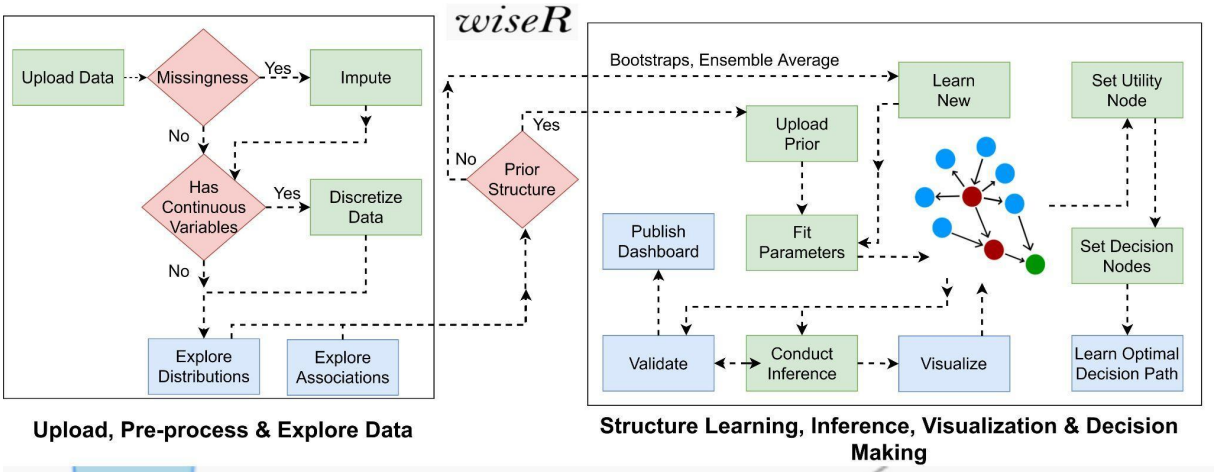
2 Implementation

wiseR was created with core R, libraries and in-house pipelines for statistical evaluation and visualization. The app is primarily built using shiny and other associated shiny libraries for the UI needs, visNetwork for the graphical needs and bnlearn, HydeNet for the working logic. Core python machine learning libraries such as pandas, numpy, scipy, pytorch are used for GNN structure learning needs processed through R. It is a cross-platform web app available to use on any system with the R (v>3.5.0) installed. The app has been tested on multiple operating systems using different browsers, ensuring wider availability. To accommodate complex data, file limit has been extended to 8000 MB and provision for parallel computing has been provided, results for which may vary from hardware to hardware. The app can easily be installed and launched as:

```
>library('devools')
>install_github('tavlab-iiitd/wiseR')
>wiser()
```

3 Functionalities

The pipeline has been designed for maximum exploitation of complex data, to assist in mining useful information and identify decision-making points which users can then learn optimal decision paths for. The app is divided into 6 tabs; 'App Settings', 'Data', 'Association Network', 'Bayesian Network' ,'Decision Networks 'and 'Publish your dashboard', each of which covers a functionality as suggested in the name. Through 'App Settings' users can enable parallel processing and choose no. of parallel nodes (hardware dependent). Dataset is uploaded and pre-processed to fit learning requirements in the 'Data' tab which also enables variable exploration among other features. A broad relation between the variables can be understood by building and exploring association networks through the 'Association Network' tab or the user can directly skip to 'Bayesian Network' for structure learning. The app enables structure learning using algorithms and parameters of choice, with provision for bootstrapped learning, editing and initializing learned structures and uploading externally learned graphs. 'Hill Climbing', 'Grow-Shrink' and 'DAG-GNN' are some of the graph structure learning methods available through the app. The learned structure can then be analyzed for insights using graph exploration tools at disposal. Variables of interest are used as evidence/event to build inference plots and actionable results can then be reported, all this can be done in the 'Bayesian Network' tab. To learn optimal decision paths for the actionable points the user can move to the 'Decision Network' tab which utilizes the learned structure as the decision network. The user can set the utility node along with the value for each outcome which they wish to maximize/minimize and set the decision nodes for which they want to learn the optimal policy. The app makes it possible to learn the best decision path on the utility node as well as compare it with other courses of actions. The app also has provision to download a customized dashboard as an R package, which can be used to explore and infer from the results independently and is groundbreaking in terms of ease of publishing. This feature is enabled in the 'Publish your dashboard tab'. A detailed tutorial with a walkthrough of the app layout and functionalities is also available in the 'App Tutorial' section as well as a supplementary folder(wiseR_Tutorial.pdf) along with a use-case, replicating validated results to show the utility of Bayesian networks.



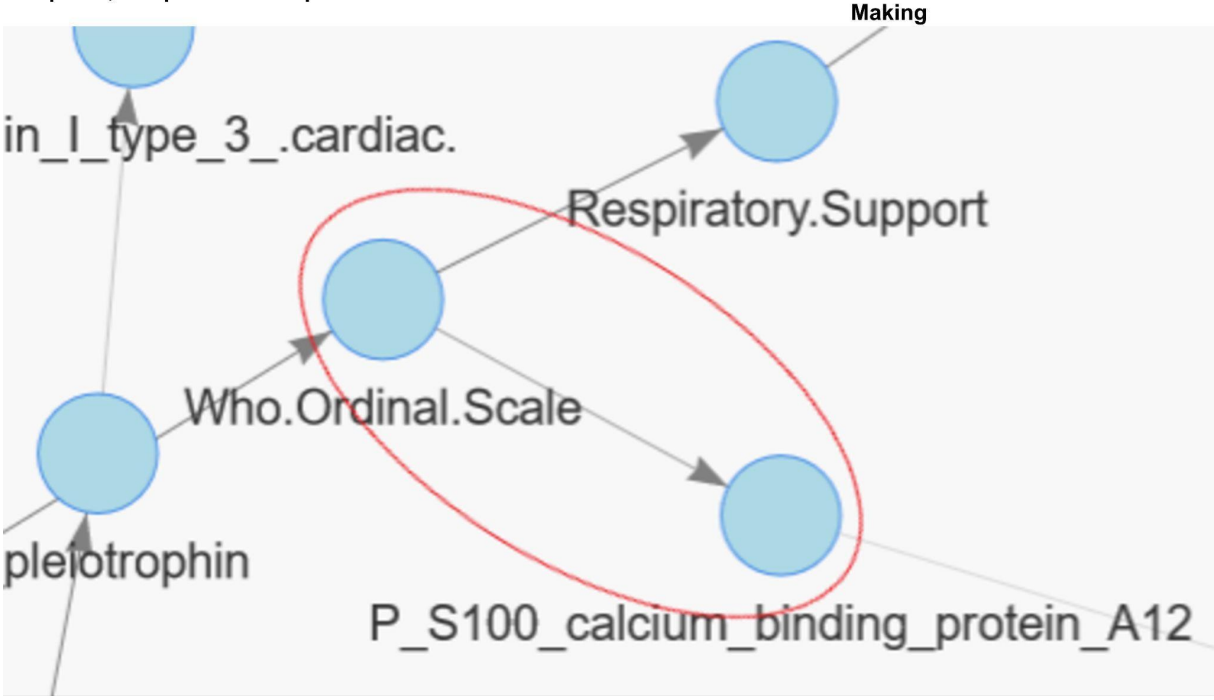


Fig. 1. Workflow of the App on the top; learned structure at the bottom verifying causality between S100A12 protein and covid-19 severity (Measures using WHO ordinal scale)

4 Case Study: COVID-19 Biomarker Discovery

Covid-19 has many countries strapped for medical resources, highlighting the need for allocation of these resources to those more in need. Several markers have been associated with severity in covid patients, none with enough prognostic power, thus need for better biomarkers is essential. We utilized the wiseR app on a comprehensive clinical and multi-omics dataset of covid-19 patients (Su et al. 2020), and discovered S100A12 as the strongest actionable node for predicting the severity of infection in covid-19 patients as described on WHO ordinal scale. The causality between the 2 is shown in the bottom half of Fig-1. As processed using wiseR. Our results were verified in a recent paper finding the potential of S100A12 as a single transcript to guide clinical decisions (Lei 2021). Interestingly, S100A12 is also a known biomarker for sepsis and wiseR was able to prioritize this from a set of potential markers, while allowing for causal discovery and probabilistic inference. A detailed walkthrough for reproducibility of results, processed dataset and exact inferences drawn from this causal link can be found in the

appendix section below. Identifying such prominent prognostic biomarkers amidst a pandemic such as COVID-19 can help translational discovery, early identification, and design of better interventions, potentially saving many lives. A dashboard to explore the network or evaluate the findings is also made available, generated through wiseR itself, highlighting its efficacy in improving reach and understanding of medical research.

5 Conclusion

We present wiseR, an end-to-end causal discovery web-application using explainable graphical models. This application has been validated and tested for numerous use cases in public health (Sethi et al. 2019; Awasthi et al. 2020; Jha et al. 2021) and we show for the first time the use of such an approach for biomarker discovery for predicting severity in COVID-19 patients. wiseR provides state-of-the-art algorithms in graph neural networks and Bayesian Networks for causal discovery. The end-to-end framework allows for deployment of these models in translational scenarios by experimental biologists and clinicians without the need for expertise in causal networks.

Acknowledgements

This work was supported by the DBT/Wellcome Trust India Alliance Fellowship IA/CPHE/14/1/501504 awarded to Tavpritesh Sethi. We are also thankful to Anant Mittal for testing the application and providing suggestions.

References

Awasthi, Raghav, Prachi Patel, Vineet Joshi, Shama Karkal, and Tavpritesh Sethi. 2020. “Learning Explainable Interventions to Mitigate HIV Transmission in Sex Workers Across Five States in India.” *NeurIPS Workshop (2020)*.

Drton, Mathias, and Marloes H. Maathuis. 2017. “Structure Learning in Graphical Modeling.” *Annual Review of Statistics and Its Application* 4 (1): 365–93.

Jha, Indra Prakash, Raghav Awasthi, Ajit Kumar, Vibhor Kumar, and Tavpritesh Sethi. 2021. “Learning the Mental Health Impact of COVID-19 in the United States With Explainable Artificial Intelligence: Observational Study.” *JMIR Mental Health* 8 (4): e25097.

Kleinberg, Samantha, and George Hripcsak. 2011. “A Review of Causal Inference for Biomedical Informatics.” *Journal of Biomedical Informatics* 44 (6): 1102–12.

Lei, Hongxing. 2021. “A Single Transcript for the Prognosis of Disease Severity in COVID-19 Patients.” *Scientific Reports* 11 (1): 12174.

Prosperi, Mattia, Yi Guo, Matt Sperrin, James S. Koopman, Jae S. Min, Xing He, Shannan Rich, Mo Wang, Iain E. Buchan, and Jiang Bian. 2020. “Causal Inference and Counterfactual Prediction in Machine Learning for Actionable Healthcare.” *Nature Machine Intelligence* 2 (7): 369–75.

Richens, Jonathan G., Ciarán M. Lee, and Saurabh Johri. 2020. “Improving the Accuracy of Medical Diagnosis with Causal Machine Learning.” *Nature Communications* 11 (1): 3923.

Scutari, Marco. 2017. “Bayesian Network Constraint-Based Structure Learning Algorithms: Parallel and Optimized Implementations in the Bnlearn R Package.” *Journal of Statistical Software*. https://doi.org/10.18637/jss.v077.i02.

Sethi, Tavpritesh, Anant Mittal, Shubham Maheshwari, and Samarth Chugh. 2019. “Learning to Address Health Inequality in the United States with a Bayesian Decision Network.” *Proceedings of the AAAI Conference on Artificial Intelligence* 33 (01): 710–17.

Su, Yapeng, Daniel Chen, Dan Yuan, Christopher Lausted, Jongchan Choi, Chengzhen L. Dai, Valentin Voillet, et al. 2020. “Multi-Omics Resolves a Sharp Disease-State Shift between Mild and Moderate COVID-19.” *Cell* 183 (6): 1479–95.e20.

Yu, Yue, Jie Chen, Tian Gao, and Mo Yu. 2019. “DAG-GNN: DAG Structure Learning with Graph Neural Networks.” *ICML (2019)*.

Appendix

Structure Learning Based Discovery of S100A12 as a marker for COVID-19 Severity

1 Result dashboard installation

> library('devools')
> install_github('tavlab-iiitd/ COVID19_Biomarkers')
> COVID19_Biomarkers::biomarkers()

2 Supplementary Material

- **data.csv**: Contains the discretized and filtered data used for structure learning and inferences
- **blacklist.csv**: Contains the list of edges blacklisted from structure learning as shown above.
- **wiseR_Tutorial**.pdf: Detailed tutorial of the wiseR application with use case examples in inference and decision from real world datasets.

3 Data Processing

- We utilize an open source covid-patient dataset. (Su et al., 2020)
- We merge the section of the data with information on Protein analysis, Metabolomes and Clinical variables. PCA is performed to trim down Protein and Metabolome variables, to keep only the most relevant and contributing ones.
- Data is then uploaded to wiseR as ".CSV" and Discretized

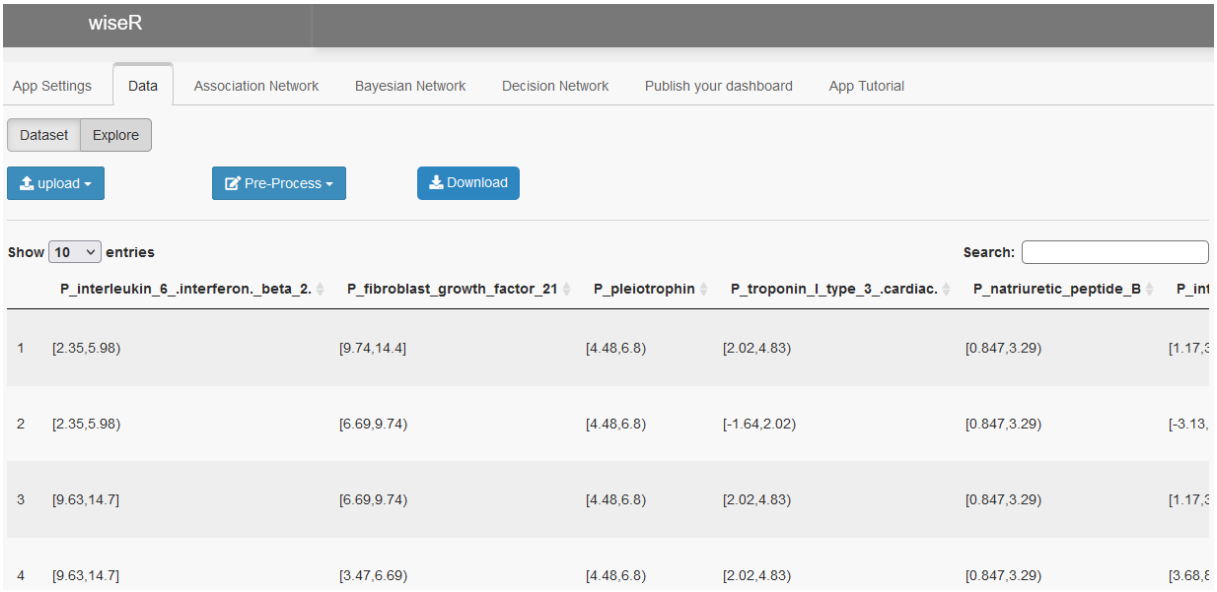


Fig. 1. Shows data uploaded and discretized through the wiseR app.

- A blacklist of edges is created to prevent flow of causality from covid variables to pre-covid variables. These blacklist edges are saved in a "from-to" table in a ".csv" file.

4 Structure Learning

- We perform Bayesian Structure Learning on the data using wiseR and the following parameters:-
    - Algorithm: Hill Climbing
    - Network Score: Akaike Information Criterion
    - Parameter Fitting: Bayesian Parameter Estimation
    - Blacklisting: Yes (csv file of blacklist provided)
    - Bootstrap Model: Yes
    - No. of Bootstrap:51
    - Edge Strength: >0.51
    - Direction Strength: >0.51
    - Sample Portion: 1

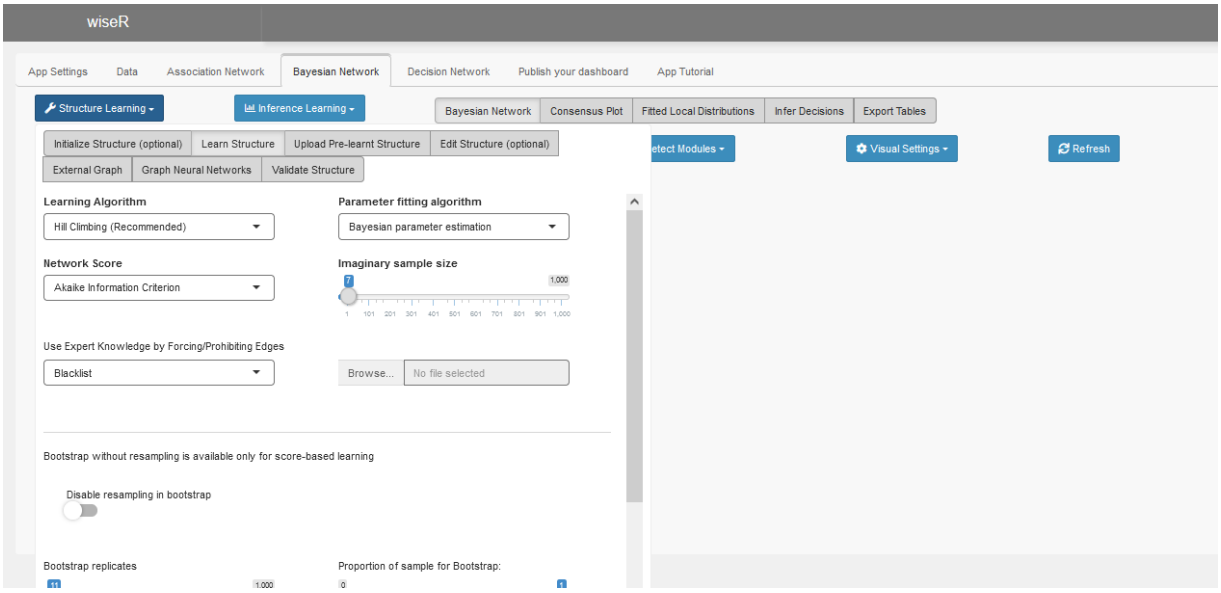


Fig. 2. Parameters of Bayesian Structure Learning

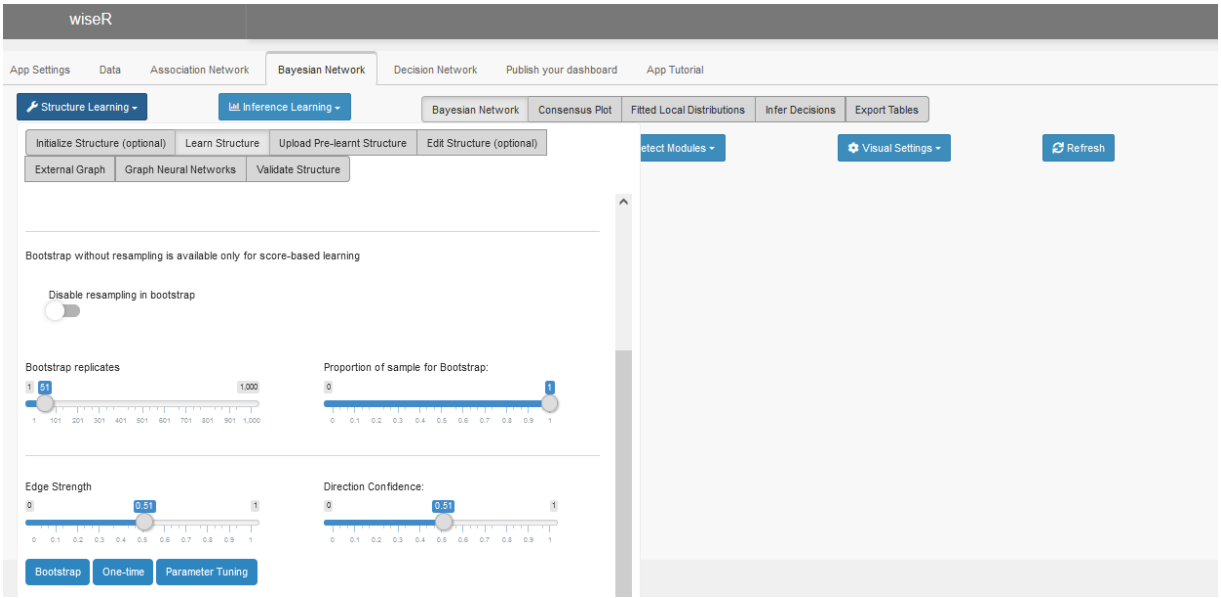


Fig. 3. Parameters of Bootstrap Learning

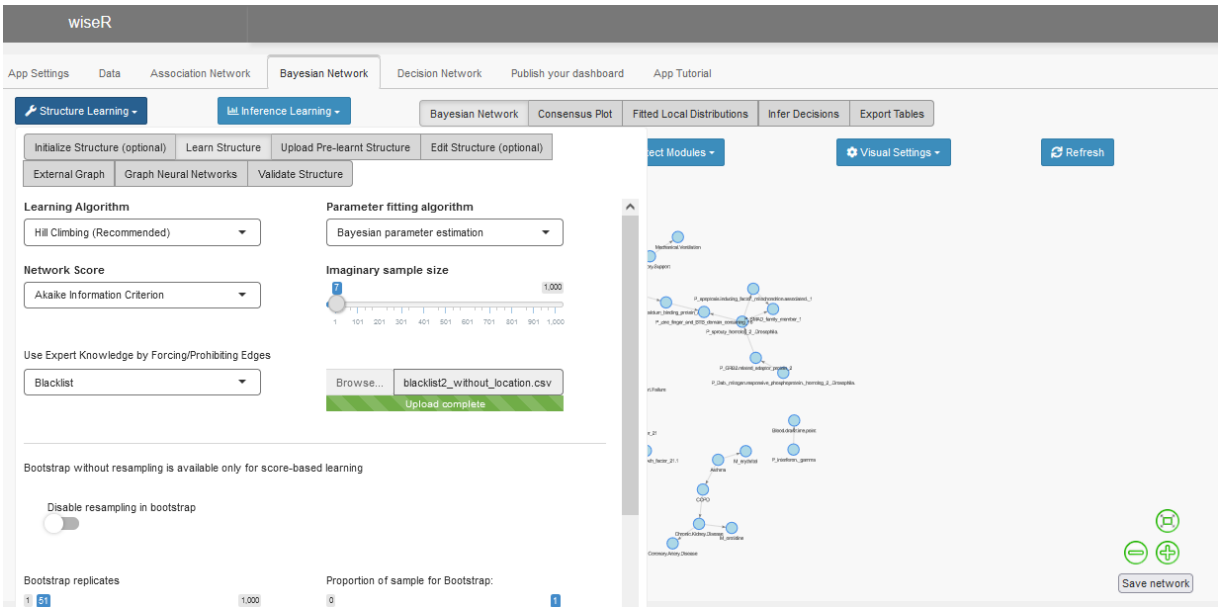


Fig. 4. Utilizing Blacklist Edges

3 Understanding Network and Inference

- We can observe in the learned structure the link between Who Ordinal Scale for Covid Severity and S100A12

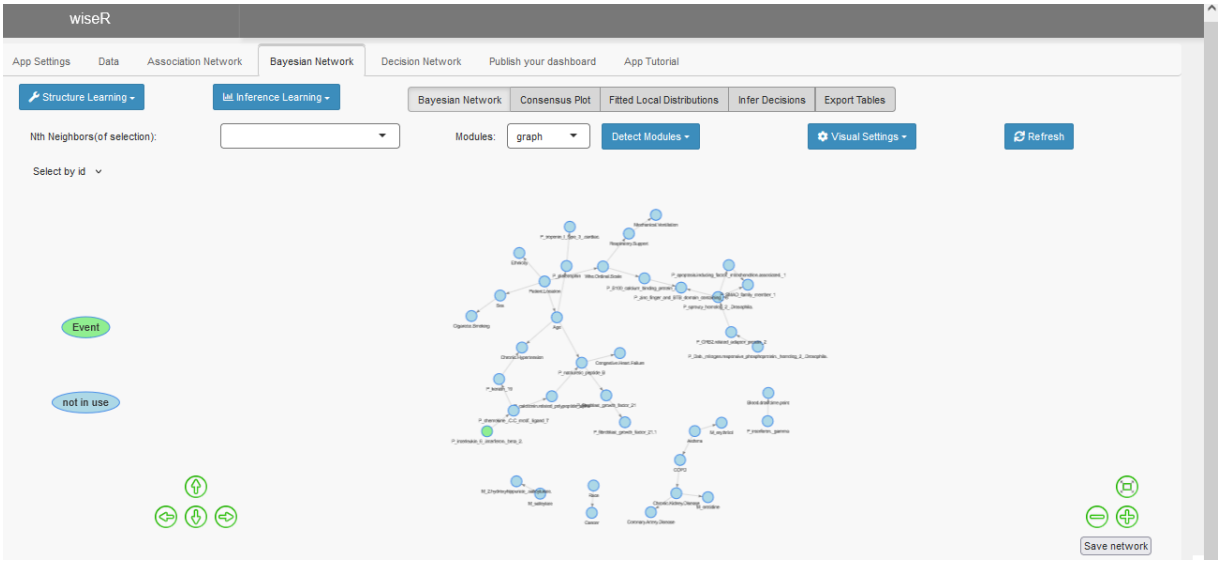


Fig. 5. Learned Network Structure

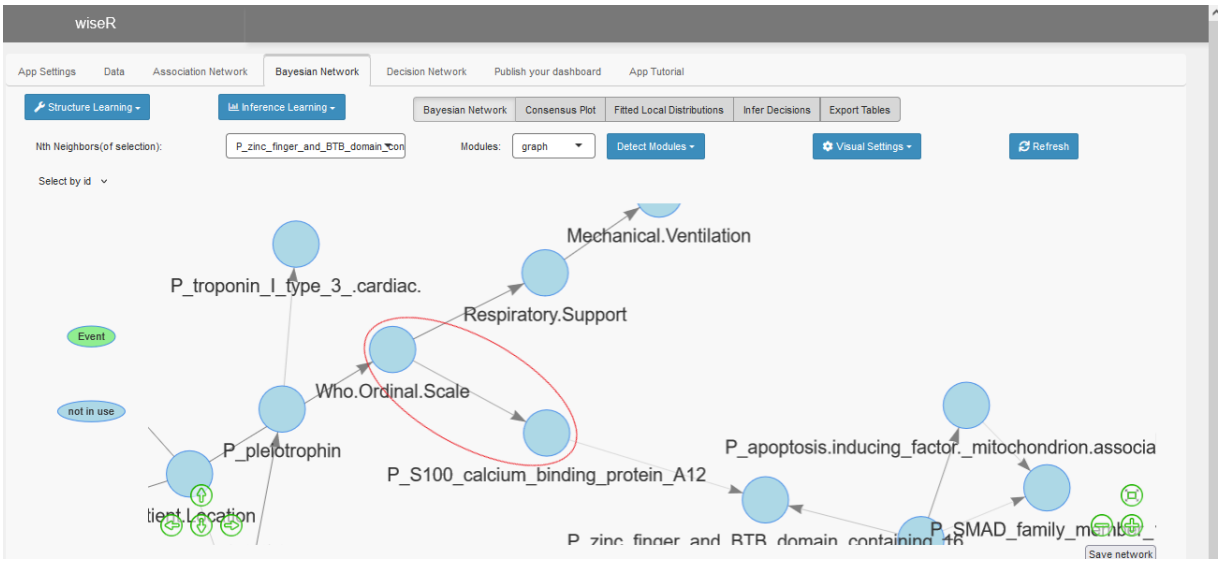


Fig. 6. Directed edge between WOS and S100A12

- Our network just validates the causality already verified (Lei, 2021)
- We use an inference plot to show that S100A12 protein markers are positively causal to covid severity. WOS rate covid severity on a scale of 1-7, 1 being the mildest to 7 being most severe. Below plots show how increased value of S100A12 protein markers indicate the likelihood of covid severity in patients.

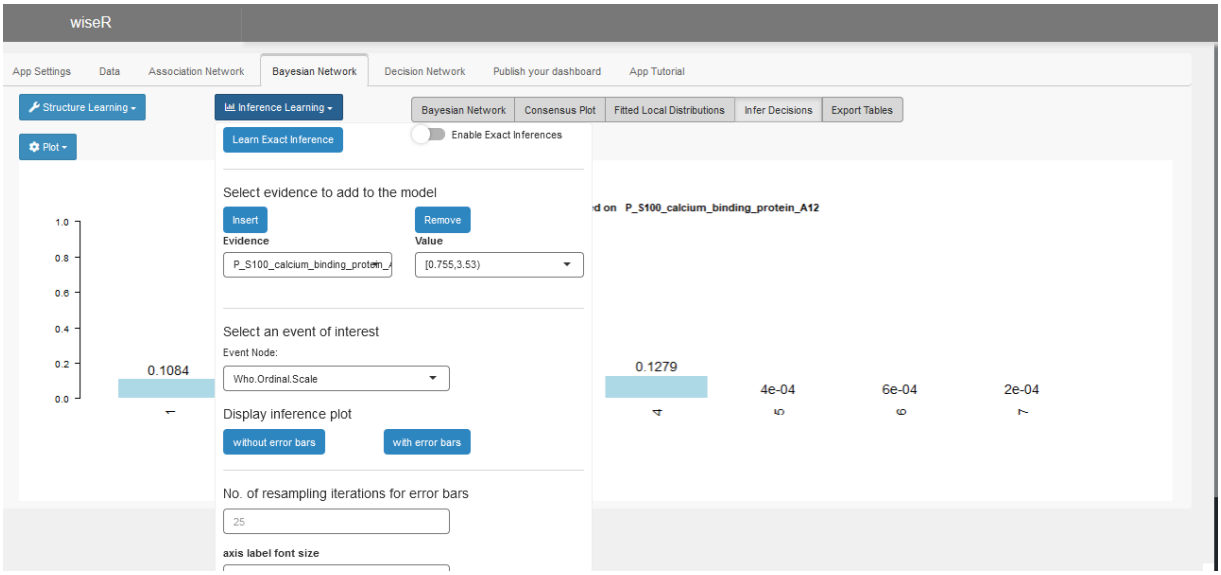


Fig. 7. Set S100A12 in the lowest range

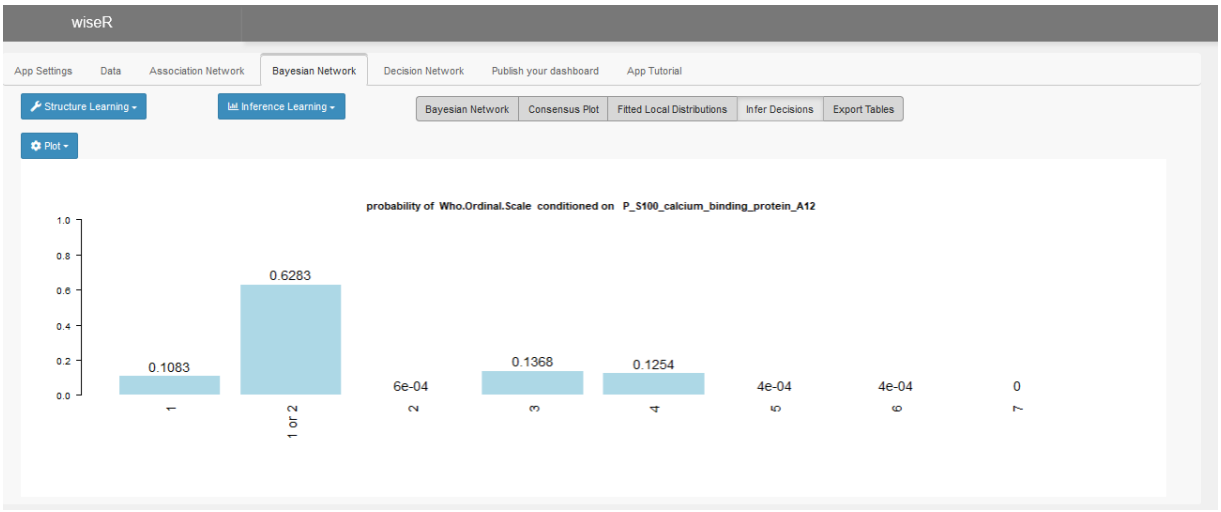


Fig. 8. Lowest Bracket for S100A12 markers indicate high chances of mild covid

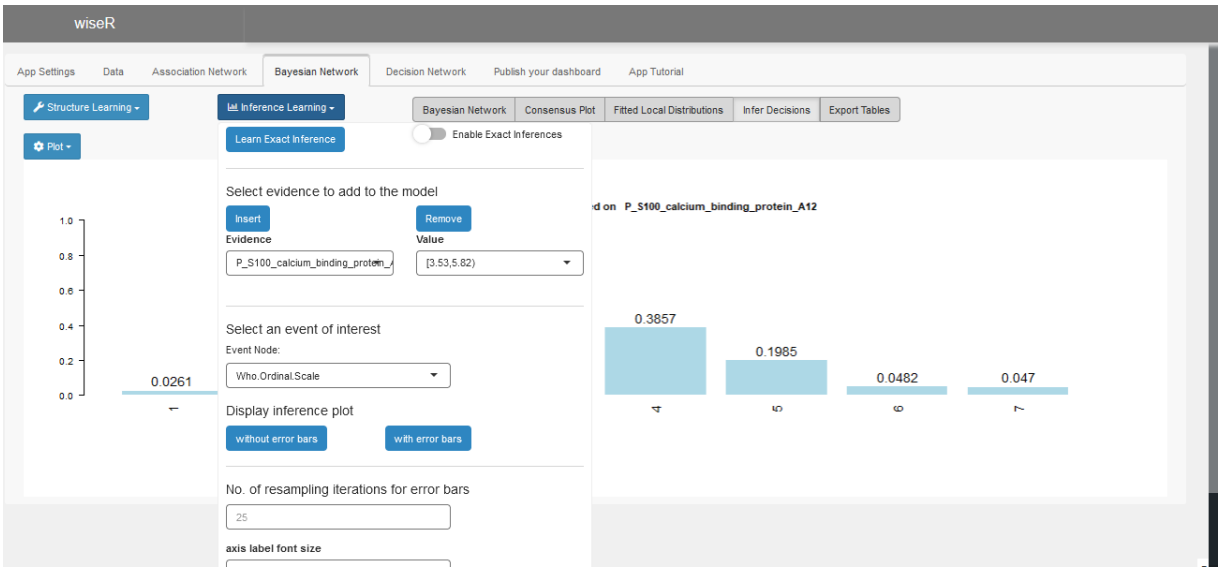


Fig. 9. Set S100A12 in the middle range

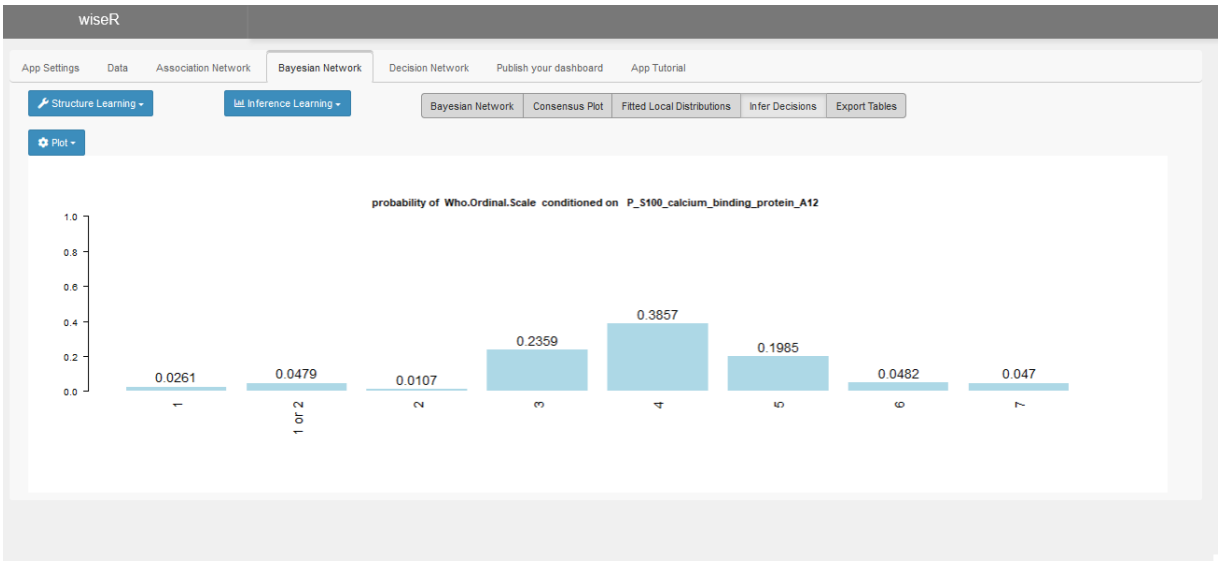


Fig. 10. Middle Bracket for S100A12 markers indicate high chances of moderate to severe covid

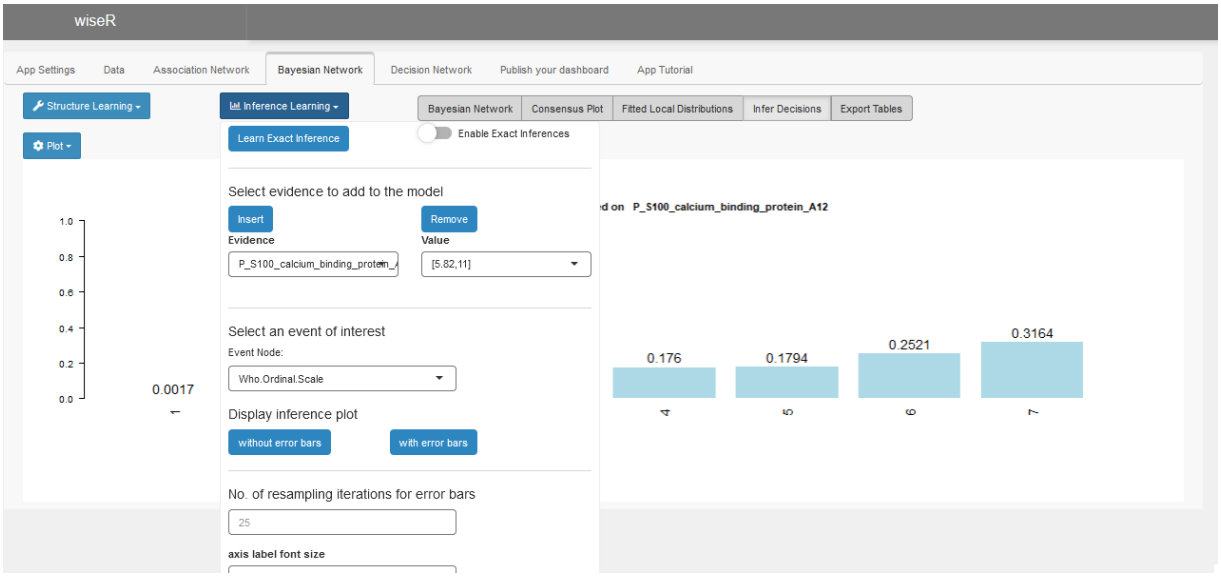


Fig. 11. Set S100A12 in the Highest range

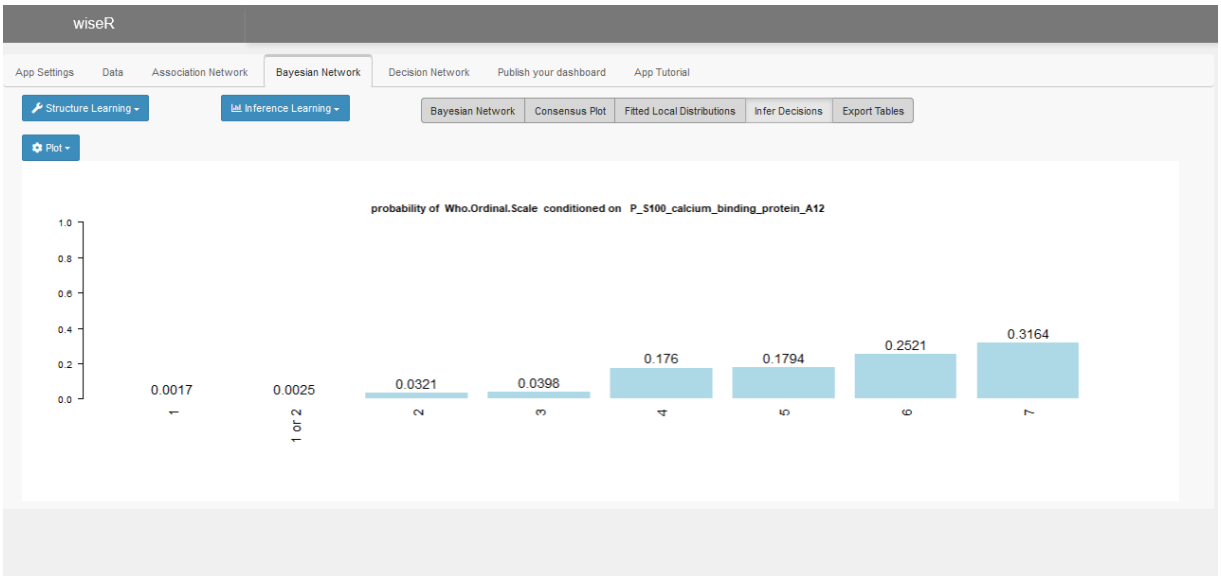


Fig. 12. Highest Bracket for S100A12 markers indicate high chances of severe covid

wiseR- Comprehensive Tutorial for Structure Learning and Decision Making

1 Getting Started

- To install latest developmental version of wiseR in R
  > **devtools::install_github('tavlab-iiitd/wiseR',build_vignettes=TRUE)**
- To launch the app:
  > **wiseR::wiser()**

2 Bayesian Networks

2.1 Why Bayesian Networks?

Networks are one of the most intuitive representations of complex data. However, most networks rely on pairwise associations which limits their use in making decisions. Bayesian Networks (BNs) are a class of probabilistic graphical models which can provide quantitative insights from data. These can be used both for probabilistic reasoning and causal inference depending upon the study design.

A BN is a directed acyclic graph and provides a single joint-multivariate fit on the data with a list of conditional independencies defining the model structure. Unlike multiple pairwise measures of association, fitting a model decreases the chance of false edges because the structure has to agree with global and local distributions. Importantly, unlike most other forms of Artificial Intelligence and Machine Learning, BNs are not a black-box model. These are transparent, interpretable and help the user in reasoning about the data generative process by looking at the motifs. This can be immensely useful in learning systems such as healthcare where feedback between the clinicians and the learning system is paramount for adoption. The structure (independencies) of a BN can be learnt directly from data using machine learning or be specified by the user, thus allowing expert knowledge to be injected. After the dependence structure within the data is learnt (or specified) the parameters are learnt on the network using one of many possible approaches. wiseR provides most of the existing approaches such as constraint based algorithms and score-based algorithms for parametrization of the Bayesian Network. Recommendations as per the state-of-the-art in the literature are specified at each step of the BN learning process (i.e. the recommendation of score based algorithms over constraint based algorithms for learning structure and Bayesian Information Criteria over Akaike Information Criteria for evaluating the fit). Parametrization of the network enables predictions and inferences. These inferences can be purely observational ("seeing" the state of a variable and its effect on neighbours) or causal ("doing" something to a variable, i.e. interventions and observing the effect on downstream nodes). wiseR provides scoring methods both for observational (e.g. Bayesian Information Criterion) and interventional (e.g. modified Bayesian Dirichlet Equivalent score) datasets.

The flexibility provided by BNs in probabilistic reasoning and causal inference has made these one of the most widely used artificial intelligence methods in Computer Science. However, the sophistication required to code all the features has limited their use by domain experts outside of computer science, e.g. clinicians. wiseR plugs this gap by creating a GUI for the domain experts and is an end-to-end solution for learning, decision making and deploying decision tools as dashboards, all from the wiseR platform.

2.2 What do the motifs (junctions) reveal about the data-generating process

Carefully learnt BNs are representations of the generative process that produced the data. These generative processes are captured in the the form of three basic motifs present in the network: chain, fork and collider junctions.

2.3 Chain (Mediator effect)

This motif is the simplest structure in a BN with a sequence of nodes pointing in the same direction. The intermediate nodes are called mediators and conditioning on any of the mediators makes the flanking nodes independent of each other. A mediator can be thought of as a mechanism, hence capturing the information about the mechanism removes the need for capturing the triggering process, hence making the triggering process independent of the outcome.

2.4 Fork (Confounders effect)

The second type of structure observed in a BN is a common parent pointing towards two or more child nodes. The parent represents the common cause and not accounting for the parent is a common source of error (confounding) in many statistical models built upon real world data. A folk example of such an effect is Age as a confounder (common cause) of IQ and shoe-size in children. Failure to condition upon age will lead to a spurious correlation between the IQ and shoe-size and change our reasoning (hence predictions) about the system being modeled.

2.5 Collider or a V-structure (Inter-causal reasoning)

These are one of the most interesting motifs in a Bayesian Network and are indicated by two nodes pointing towards a single child node. Although the two parent nodes are not causally connected, conditioning on the state of the common child opens up a ghost-path for probabilistic inference flow (inter-causal reasoning) which can explain many intriguing paradoxes (e.g. the Monty Hall problem).

3 Walkthrough of wiseR functionalities

3.1 Pipeline

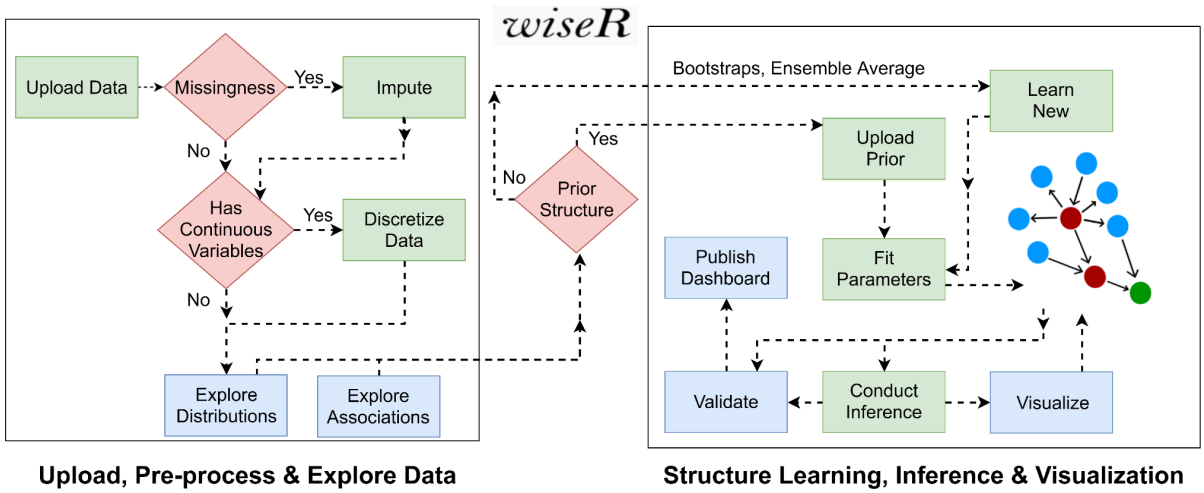


Figure S1. Flowchart showing logic of wiseR

3.2 Start

Calling the wiseR() function from the wiseR package launches the application in the default browser. The recommended browser for wiseR is Chrome. On launching the application for the first time, it conducts background checks, which takes 5-10 seconds depending upon the machine. This delay only happens on the first launch and all subsequent launches take less than a couple of seconds to initialize the user interface.

The toolbar on the left side of the page has icons for navigating to home, analysis engine, github development version and team information at any point of time.

Click “Start Analyzing” for navigating to the analysis engine with tabs specifying the functions. Hovering over most of the tabs brings up a tool-tip that briefly describes the function performed on that action.

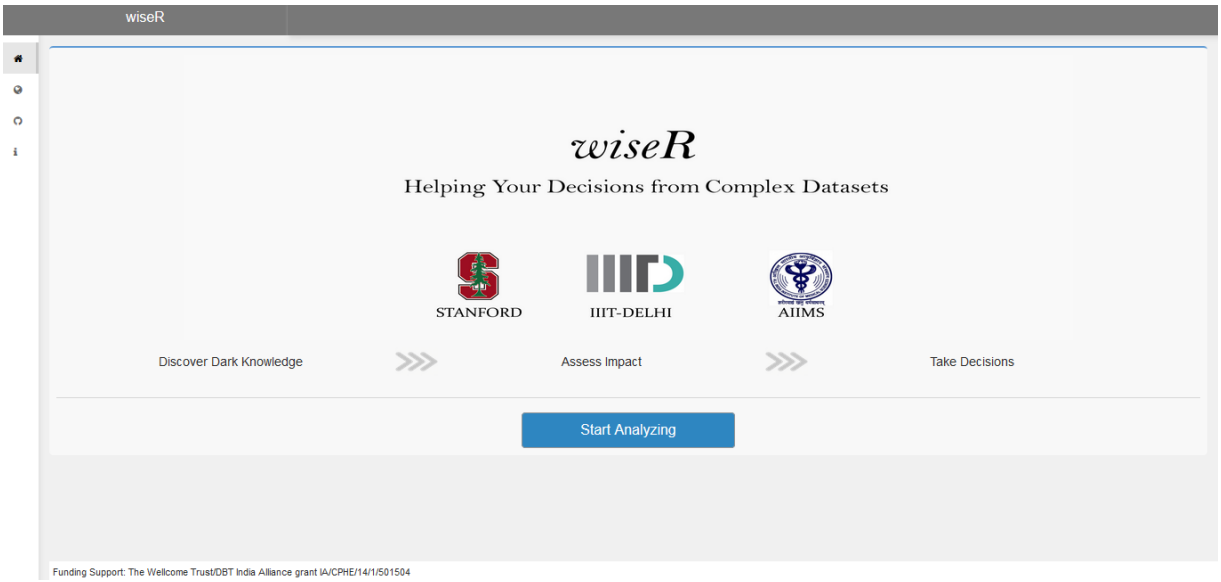


Figure S2. Homepage of the wiseR application

3.3 Main page of the wiseR engine

“Start Analyzing” takes us to the main page of the wiseR engine. An intuitive left-to-right ordering of tabs guides the user into the analysis. This page has 5 tabs named ‘App Settings’,‘Data’,‘Association Network’,‘Bayesian Network’ and ‘Publish your dashboard’, each tab covering a specific functionality (Figure S3). Each of these is described next.



Figure S3. Functional tabs on the main page of the wiser engine and the first tab (App Settings) for parallel computation

3.4 App Settings

This tab (Figure S3) is used to set the parallel computing option (number of cores) for learning BN structure. Learning structure is known to be an NP-hard problem and may be time consuming on large datasets. Setting the number of clusters allows each core to learn a structure when bootstrap learning is performed. Since learning one structure cannot be parallelized, this is an example of task parallelization.

3.5 Data

This tab is used to load, pre-process and explore the dataset (Figure S4.)

The dataset panel contains the upload and preprocess menu while the explore panel is used to visualize the distribution of the data.

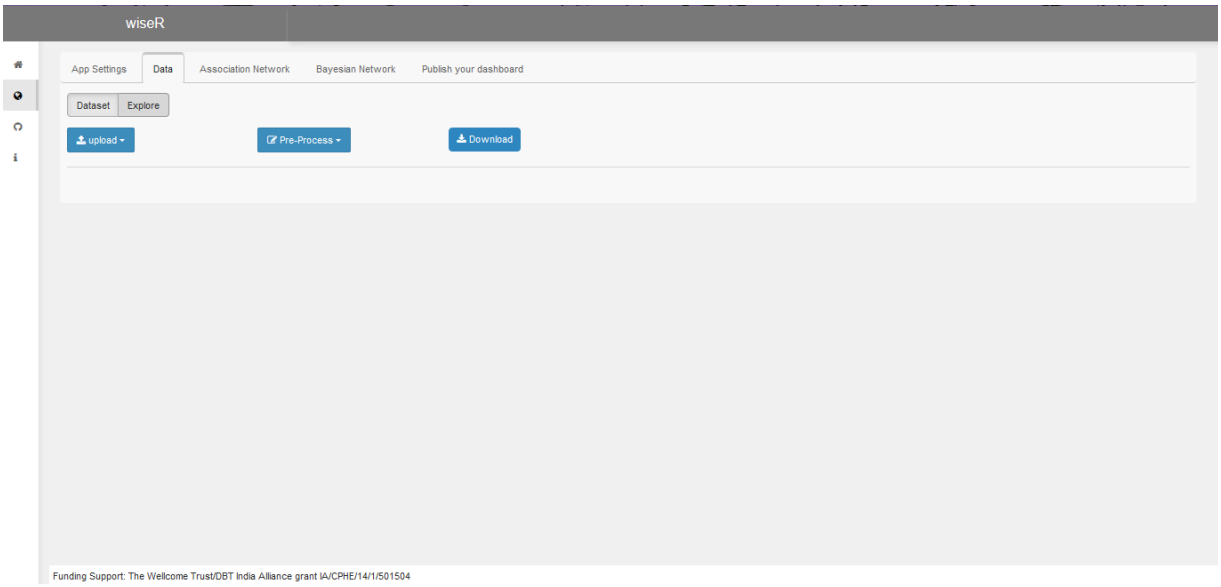


Figure S4. Data upload, exploration and pre-processing functions.

3.5.1 Upload

Users can upload their own data in one of the '.CSV','.RData' & '.txt' formats. For large data, saving the data as .RData is recommended for its compression of data, hence efficient storage. For '.txt' files different options of file formatting are provided such as:

- Comma Separated
- Semi-colon Seperated
- Tab separated
- Space Separated

New users trying to learn about BNs or testing the functionality of the app can work with the preloaded datasets (Alarm, Hailfinder, Asia, Coronary) to replicate BN analysis commonly demonstrated with these datasets.

Users must note that the default setting of the app is to treat all variables with less than 53 levels as factor variables. Variables with more than 53 levels are treated as numeric variables and the user is prompted for discretization of such numeric variables to factors. While this is a convenient general option, it may be true that a numeric variable may have less than 53 levels in a particular dataset. In such cases, the user can un-check this option and specify the numeric variables explicitly to be converted to factors in the next step.

Current limit of file upload is 8000 mb, however it is recommended to use .RData for large datasets.

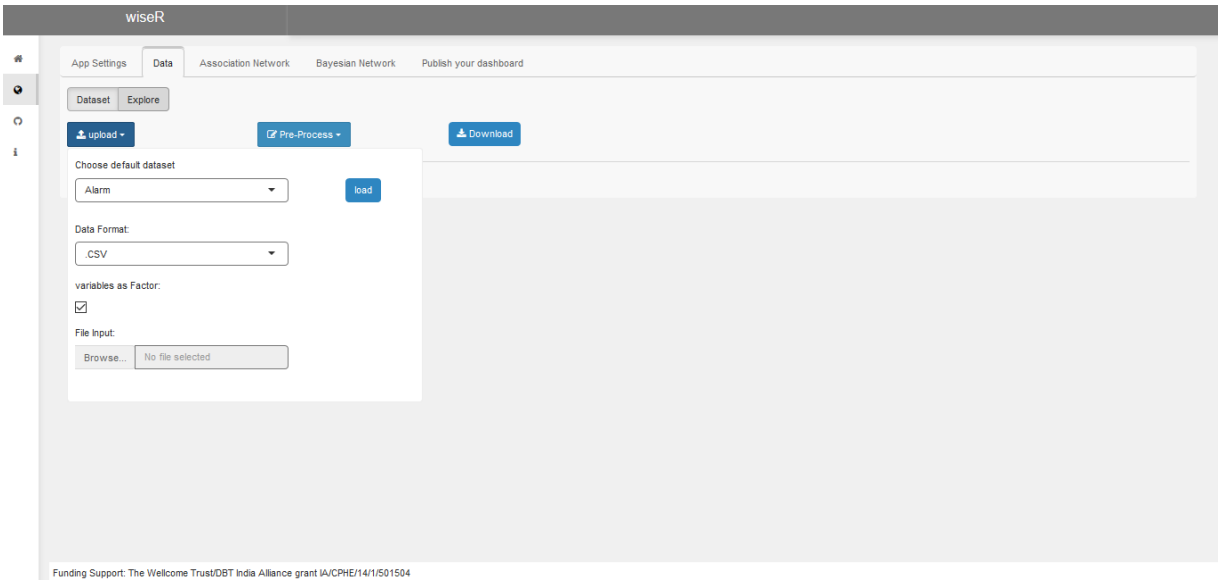


Figure S5. Features of the upload subsection of the Data tab.

3.5.2 Pre-process

The preprocess menu is used to edit the data before using it through the app. It is essential that data is discrete and complete for learning association networks and Bayesian networks. Multiple options to pre-process the data have been provided.

- User can convert any variable which has been misclassified as numeric to factor
- User can convert any variable which has been misclassified as factor to numeric
- User can choose from a number of algorithms to discretize their data - Hybrid(Recommended),Hartemink(Recommended),K-means,Quantile,Frequency,Interval
- Hybrid discretization is the default and it automatically chooses the best option for discretizing the variable. If discretization fails with the current method, it uses the next preferable method in the following sequence- k-means, quantile and interval discretization in that order of preference.
- Hartemink requires 2 parameters 'break' and 'ibreak' which can be passed using the text input boxes.
- Transpose the data frame if necessary.
- Sort variables alphabetically.
- Drop/Reset variables from the analysis. Using the delete button after entering variable names drops these from the analysis and reset button restores the original data.them from the data using the delete button.
- Select Interventions. The learning algorithm will recognize the selected variable as intervention, and will use modified Bayesian Dirichlet Equivalent score for structure calculation after accounting for the intervention.

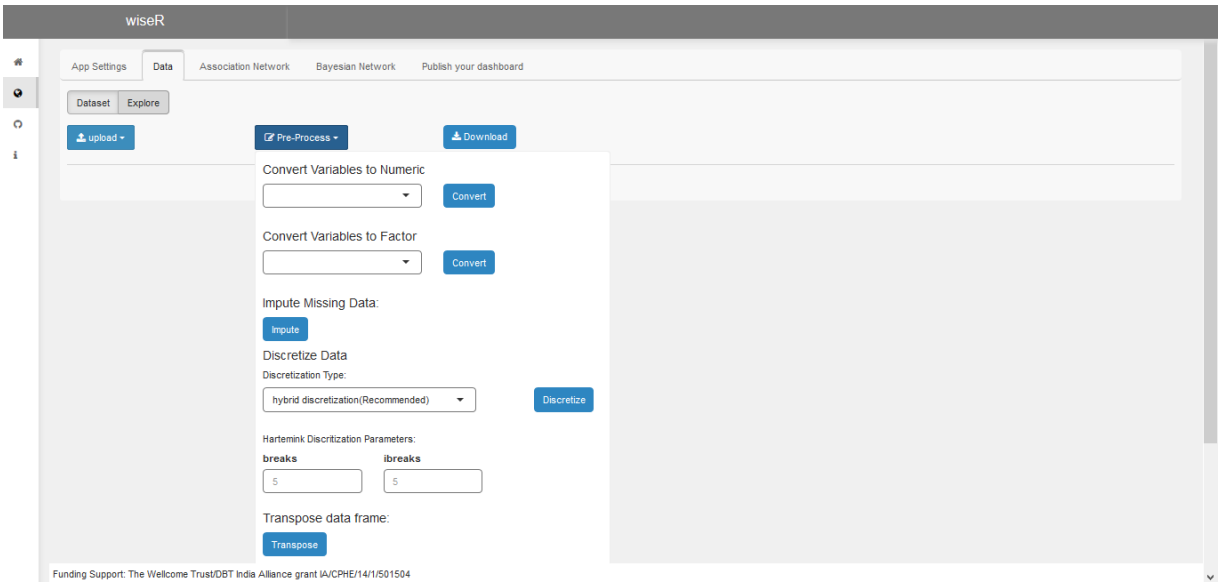


Figure S6. The data pre-processing sub-section of Data tab allows imputation, discretization, type-conversion, variable deletion and transposing and setting of interventional variables.

3.5.3 Download

Users can download the pre-processed dataset in current state as a .CSV file

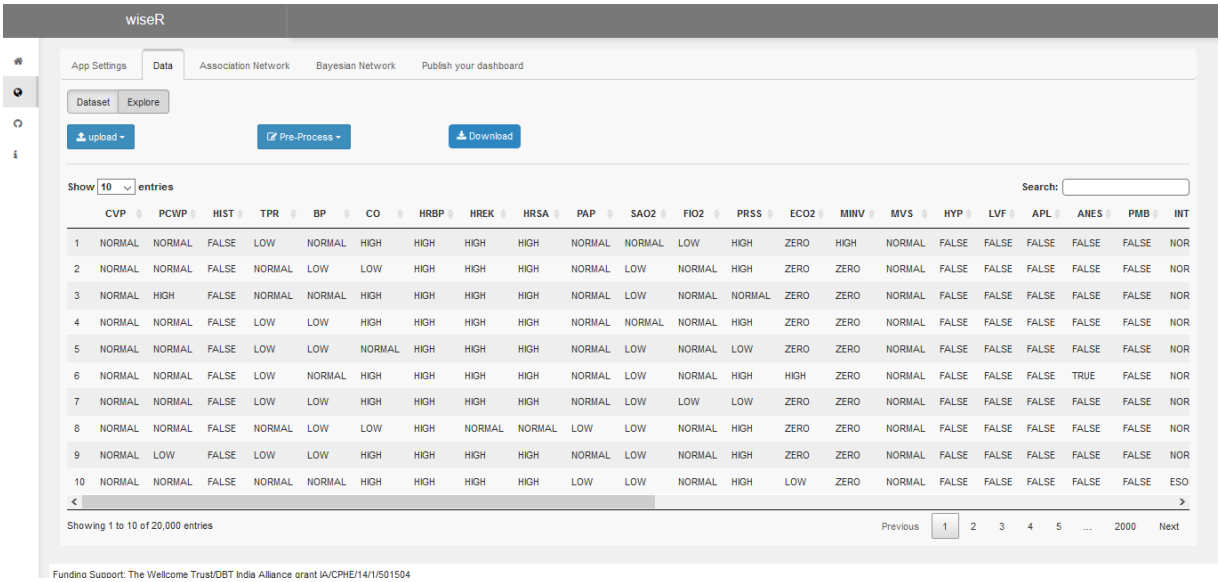


Figure S7. Users can download the pre-processed data as CSV files.

3.5.4 Exploratory Data Analysis

Pre-processed data can be explored by looking at their frequency distributions visualized as barplots. This is an important step in the analysis as it allows the users to catch any anomalies that may bias the downstream analysis.

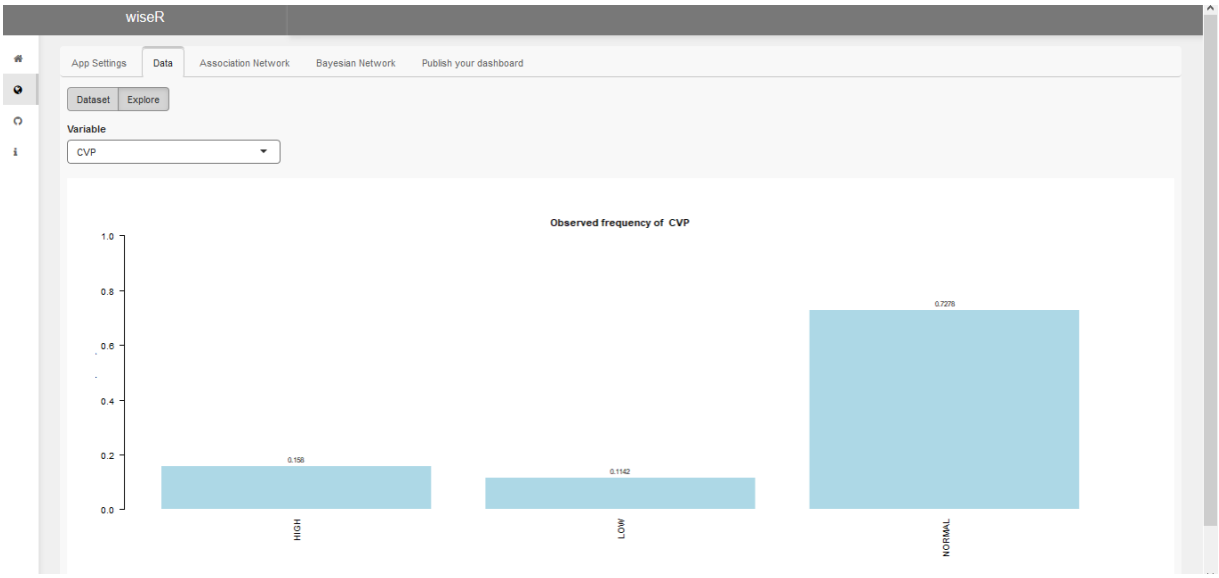


Figure S8. Exploratory data analysis such as barplots of distributions.

3.6 Association Network

Before going to Bayesian Network analysis, it is instructive to visualize the association graph constructed upon the data. Learning association networks on the data before Bayesian learning can give useful insights. Additionally the association network can serve as a starting structure for Bayesian Network structure learning. This can be achieved by using the edgelist downloaded from association network analysis in the structure learning part of the app. Various association scores are provided in the app as explained in the next section.

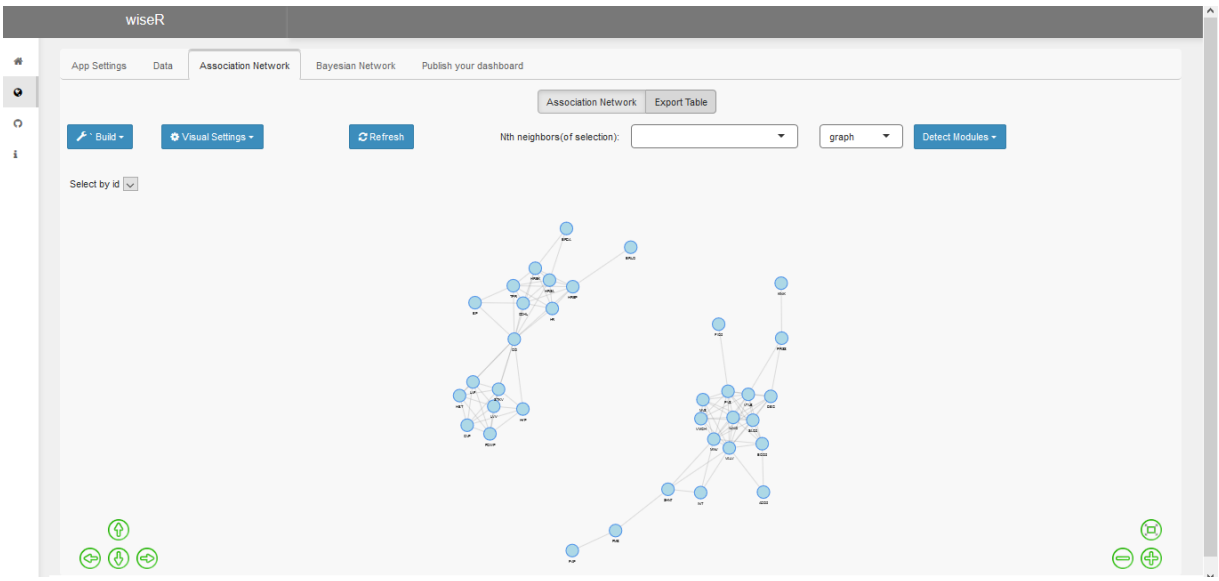


Figure S9. Build and Explore Association networks

3.6.1 Build

The build menu is used to build association networks. It consists of different metrics for association strength which are used to build association networks including:

- cramer's V(Recommended)
- cohen's D
- Goodman kruskal lambda
- Tschuprow's T

Each method returns a value between 0 and 1. 0 being two variables are not associated and 1 being highly associated. Users can select a threshold value between 0 and 1 for eliminating edges below the threshold.

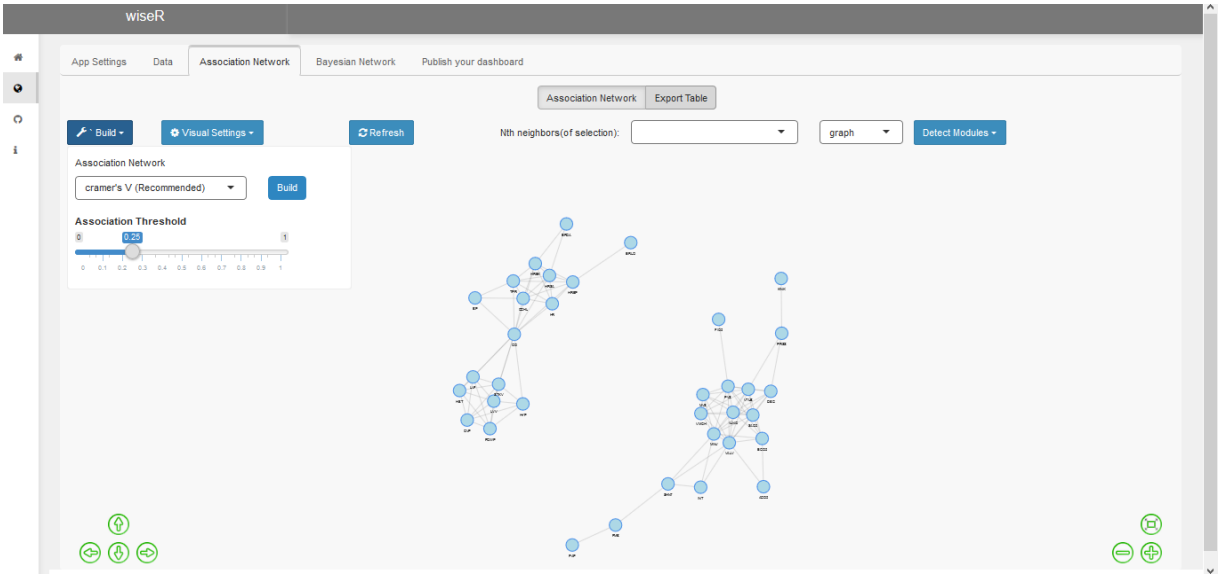


Figure S10. Options to build association networks

3.6.2 Detect Modules

Most real world networks have communities. These may form due to preferential attachment to hub nodes, often leading to a scale free network structure. To leverage this phenomenon, the app has a module/community detection pipeline revealing sub-graphs in the original graph. These graphs can be individually explored and analysed from the drop-down menu.

There are different clustering algorithms available in the app for module/community detection which is powered by linkcomm package, like ward.D (Recommended), ward.D2, single, complete, average, mcquitty, median, centroid for users to choose from.

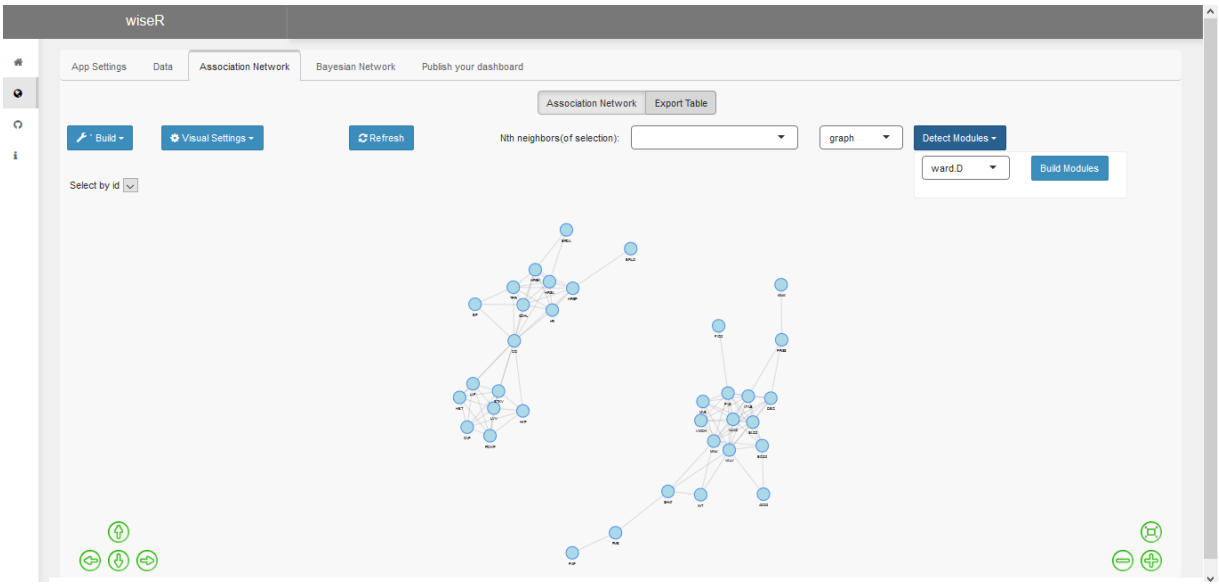


Figure S11. Detection of communities within the association network.

3.6.3 Visual settings

Visualization is often key in deriving useful inferences from a complex dataset and these functions are provided through the "Visual Settings" menu.

- **Color/Shape attributes:** The user can either select variable names or provide a vector of variable indices that need to be grouped for node shape and color. These tools are useful in analyzing bipartite or multi-partite networks.

- **n-Nearest Neighbor Visualization:**To ease the exploration of large graphs a threshold for visible neighbor chain is provided, such that clicking on a node only displays its neighbors upto that degree. Alternatively, a user can see the list of n-th degree neighbors by clicking on a variable return. This returns a list of nth degree neighbors. n can be interactively set.

- **Graph layout:** Further, the app allows various graph layout options, some of which are more suitable for particular kinds of data. Useful layouts include tree (for directed), Sugiyama (for directed), star and circle (both for undirected graphs).

- **Font Size:** User can adjust the font size of the node labels to improve visibility

- **Download graph:**Users also have the option to download the network graph in html format from the current tab or using the save network option present in the bottom right corner of the graph.

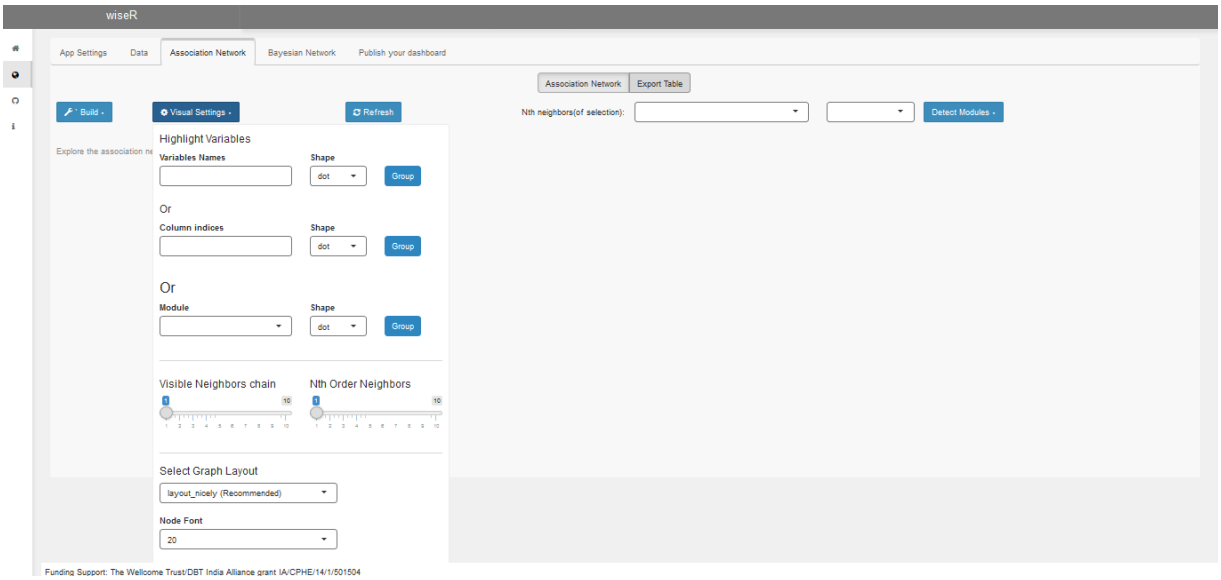


Figure S12. Visual settings to explore association graphs.

3.6.4 Export tables

Association networks are essentially a list of edges with corresponding weights. This data obtained from association networks analysis can be exported as an edgelist in tabular form and can also be downloaded as a .CSV file to allow users to switch between different networks analysis tools that can utilize edgelists.

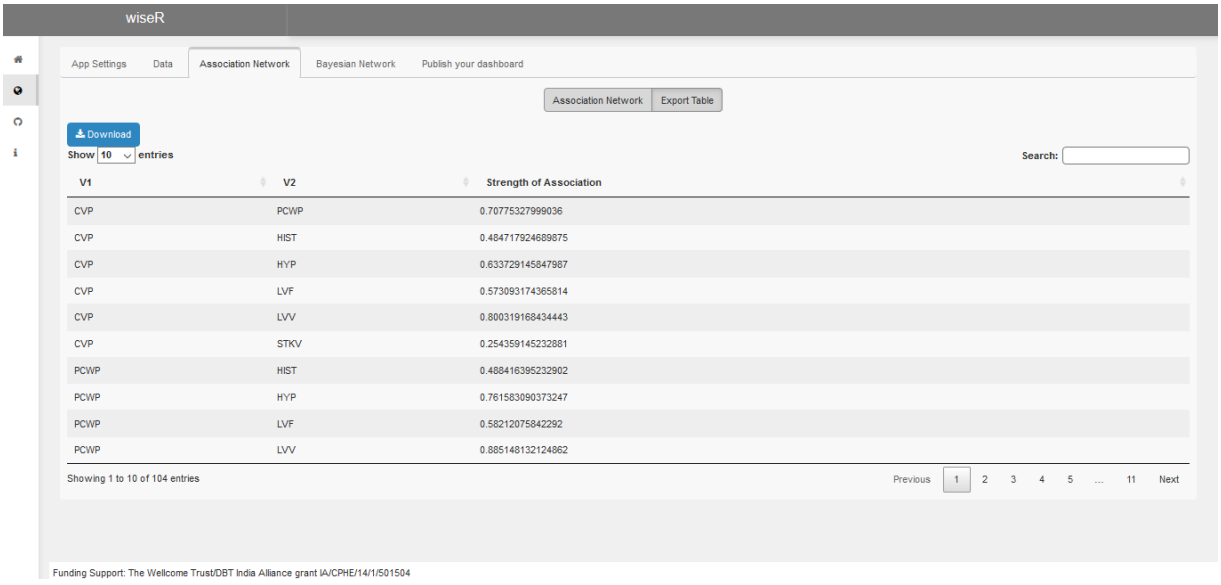


Figure S13. Export the association graph

3.7 Bayesian Network Analysis

This is the core functionality of the wiseR app. Multiple features have been added to state-of-the-art practices including bootstrapping, ensemble averaging, error bars of approximate inference etc. These ensure robustness of reasoning and decisions derived from the BN.

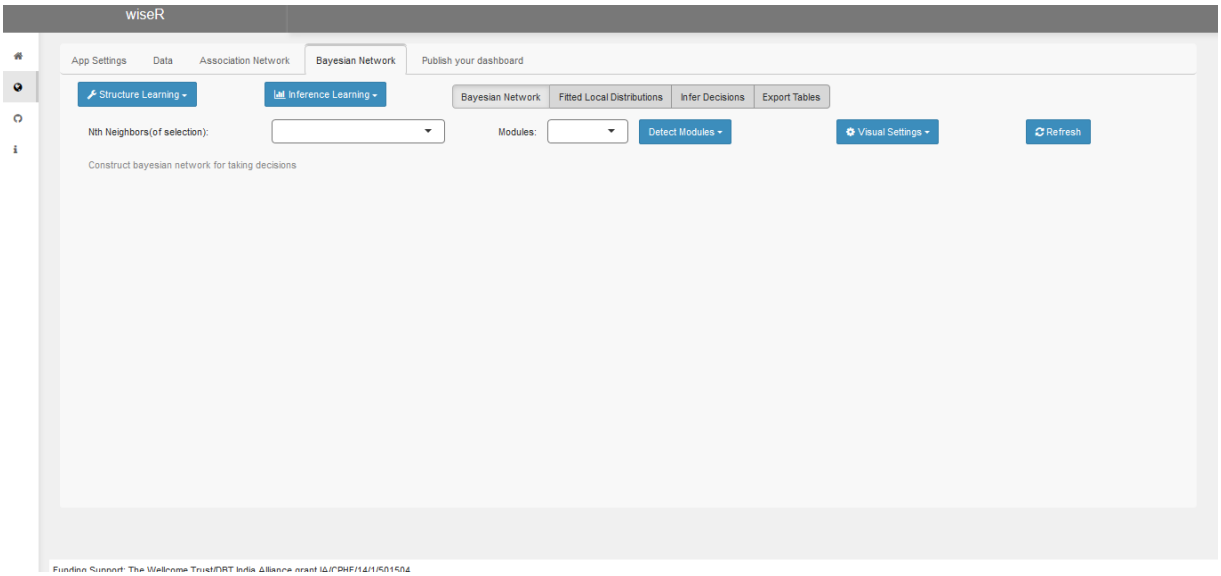


Figure S14. Build and Explore Bayesian Networks

3.7.1 Structure learning

The structure of a graphical model reveals the generative process of the data at hand. For this reason, BNs are interpretable and intuitive. Learning the structure from data is the most challenging yet insightful step in a BN analysis. However, if the structure is already known, or can be guessed, this can be pre-specified by an expert. In this way, BNs represent a perfect hybrid of expert and data-driven learning.

3.7.1.1 Initialize structure

This section relies on expert/prior knowledge and enables the user to upload a network graph as a .CSV file to initialize the bayesian learning graph. The user can also add, remove or reverse the direction of any arc, once the graph has been uploaded. (Only used in score-based learning)

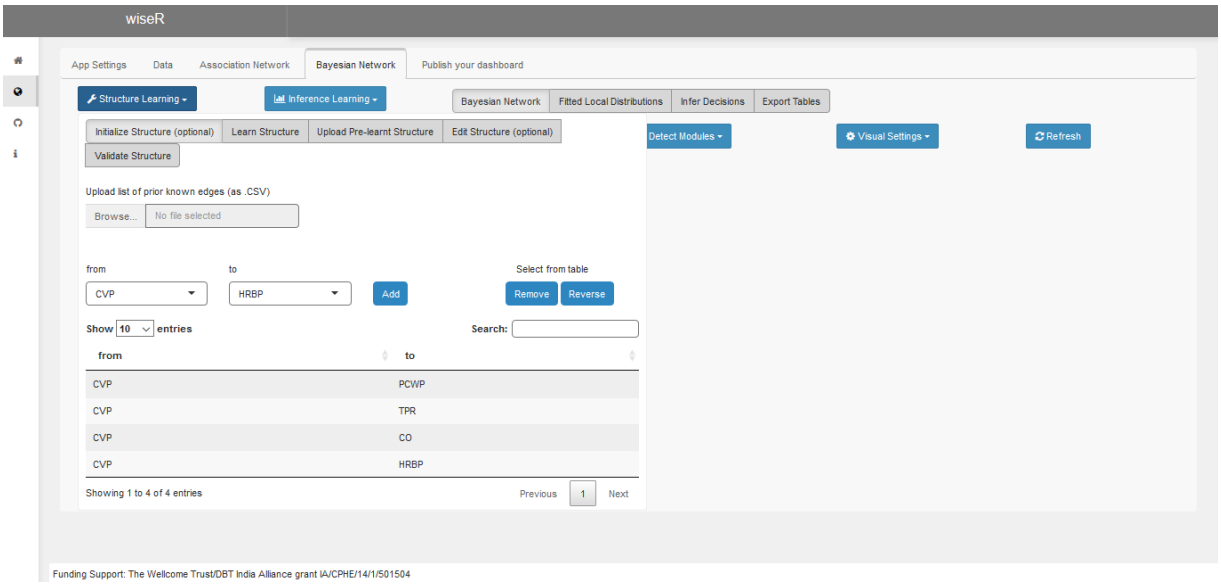


Figure S 15. Graph initialization settings

3.7.1.2 Learn Structure

In the absence of expert knowledge, independencies can be directly learnt from the data itself. The computational learning of structure can be done using a number of bayesian learning algorithms such as

- Score-based (Recommended):- Hill Climbing, Tabu
- Constraint-based:- Grow-Shrink, Incremental Association, Fast IAMB, Inter IAMB,PC
- Hybrid Learning:- Max-Min Hill Climbing, 2-phase Restricted Maximization
- Local Discovery:- Max-Min Parents and Children,Semi-Interleaved HITON-PC, ARACNE, Chow-Liu

Options used in structure learning include:-

- Parameter fitting algorithm:- Bayesian parameter estimation (recommended), maximum likelihood parameter estimation
- Network score:- Bayesian Information Criterion, Bayesian Dirichlet Equivalent, modified Bayesian Dirichlet Equivalent, log-likelihood,Akaike Information Criterion,Bayesian Dirichlet Sparse,Locally Averaged Bayesian Dirichlet (Only used in score-based learning)
- Imaginary sample size (Only used in score-based learning)
- Incorporate expert knowledge via blacklisting (explicitly restrict edges in structure learning) and whitelisting (explicitly enforce edges in structure learning) edges by uploading a .CSV file for the same
- Disabling data resampling in bootstrap learning. This is particularly useful for data with low sample size. In this case, the data remains the same but the graph is re-initialized every time. This method is only possible for score based algorithms as it requires random graph initialization.
- Bootstrap iterations:- Bootstrapping means sampling the data with replacement. This parameter specifies the number of bootstrap iterations to run for bootstrapped structure learning.
- Bootstrap sample size:- Proportion of data to be used as sample for bootstrap learning.
- Edge strength:- Set a threshold value between 0 and 1 to remove edges below specified strength from the final structure. Edge strength can be interpreted as the probability of occurrence of an edge in bootstrap learning.

- Edge direction:- Set a threshold value between 0 and 1 to remove edges below specified direction confidence from the final structure. Edge direction can be interpreted as the confidence of direction of a learned arc in bootstrap learning.
- Although not recommended, bootstrap based structure learning may be by-passed via the “Direct Learning” option for learning the structure without bootstraps. Please note that this is suitable only for quick exploration and may give less robust structure.

note* The user does not have to do the bootstrap learning every time if they wish to explore different thresholds for edge and direction strength. This can be achieved by using the parameter tuning option.
The final learnt structure can be saved as a bnlearn object in .RData format through the save option

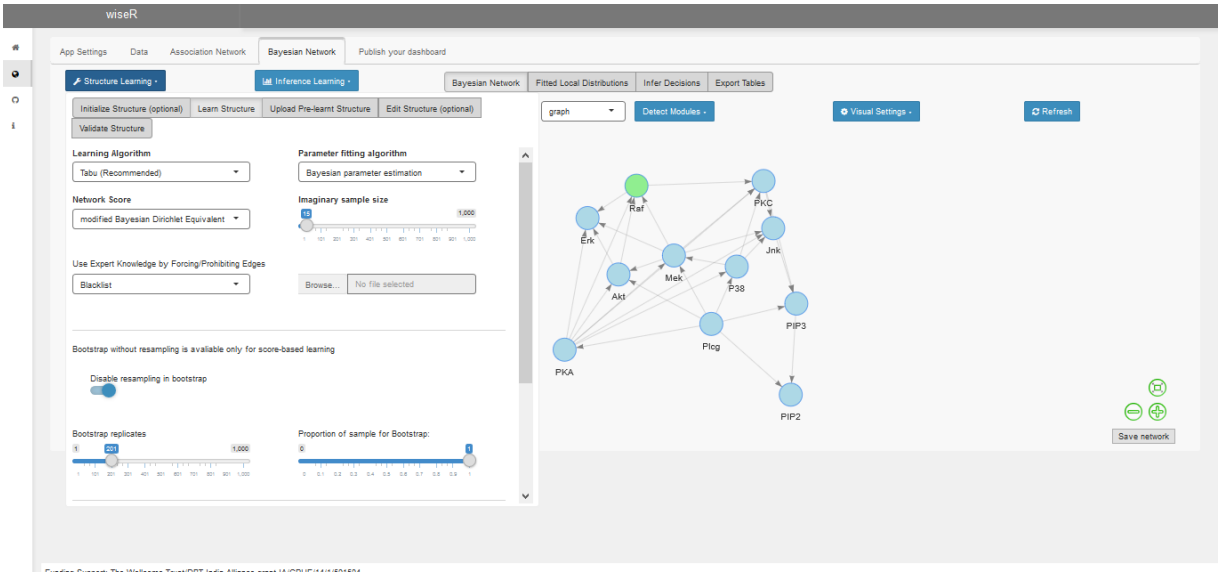


Figure S16. Structure learning settings

3.7.1.3 Upload pre-learnt structure

A user returning to analysis should not need to repeat the above steps. An already learnt and saved bayesian network structure can be uploaded as an .RData file.Options to upload both bootstrapped and direct structure and adjust their parameters are available in the app.

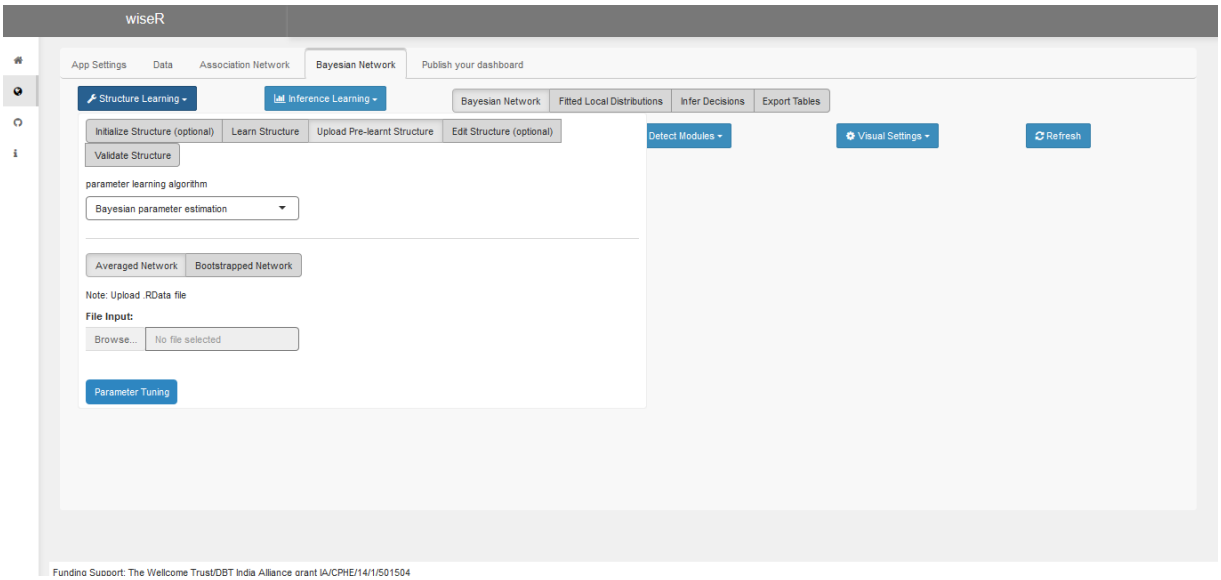


Figure S17. Upload pre-learnt structure

3.7.1.4 Edit structure

wiseR allows expert checks at every step and an expert can identify edges and directions that are not plausible in the real world (e.g. smoking cannot influence natural gender, but the reverse can happen). An expert can add/remove arcs, or reverse the direction of the existing ones on the basis of plausibility.

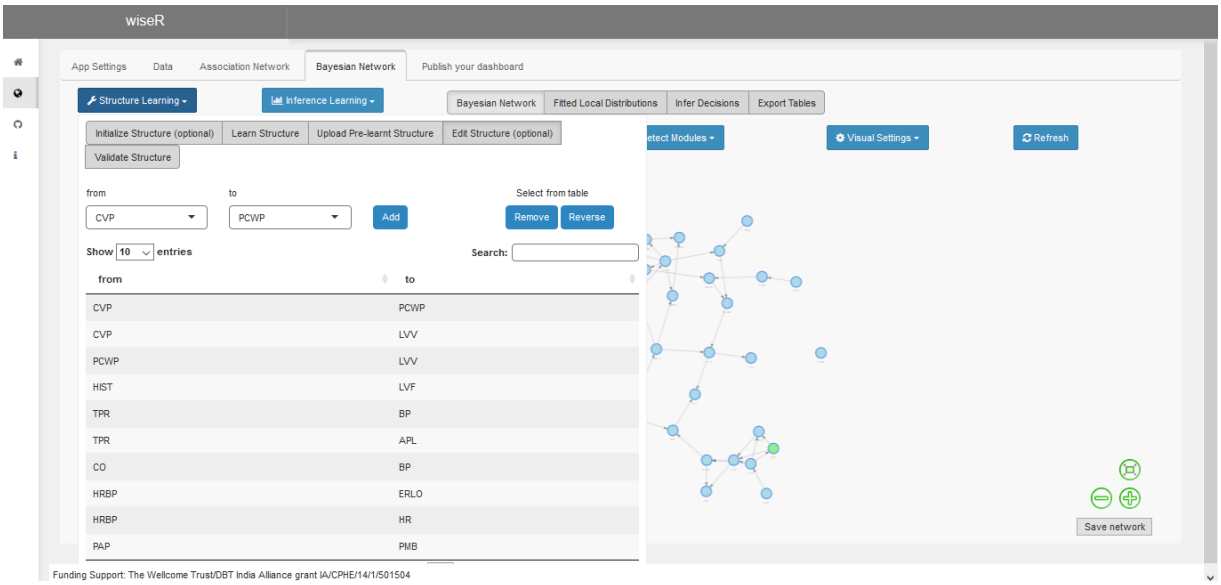


Figure S18. Edit learned structure via adding, removing or reversing directions based upon expert knowledge and plausibility.

3.7.1.5 External Graph

wiseR allows users to work with an external graph through the app. It is possible users may want to test out newer graph learning algorithms which are not available through the app. This ensures the user access to app features provided they have a edgelist of the external graph in “from” and “to” structure saved as a “.csv”

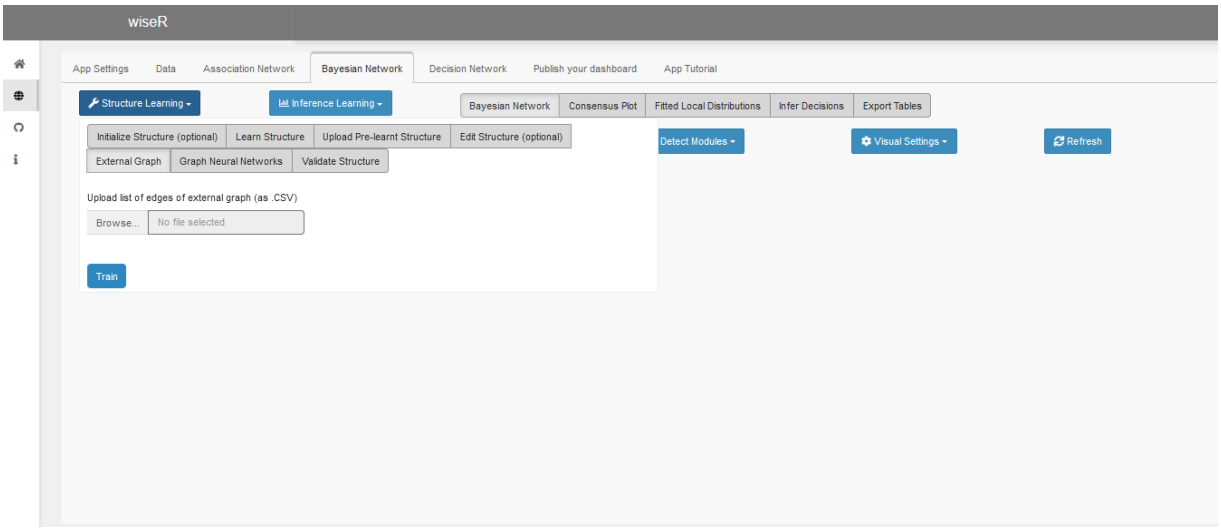


Figure S19. Provision for uploading externally learned graphs to work through the App.

3.7.1.6 Graph Neural Networks

With the current advances in Graph based machine learning, wiser provides the standard GNN based structure learning method through the app. The user is asked to provide the Python and Conda environment path with all requisite libraries such as:

- pytorch
- matplotlib
- networkx
- pandas
- scipy
- scikit-learn
- argparse
- Causalnex
- Ipython

The user can provide the no. of bootstraps for averaged structure learning.User can choose between DAG-GNN (recommended) and NoTears algorithm for structure learning.

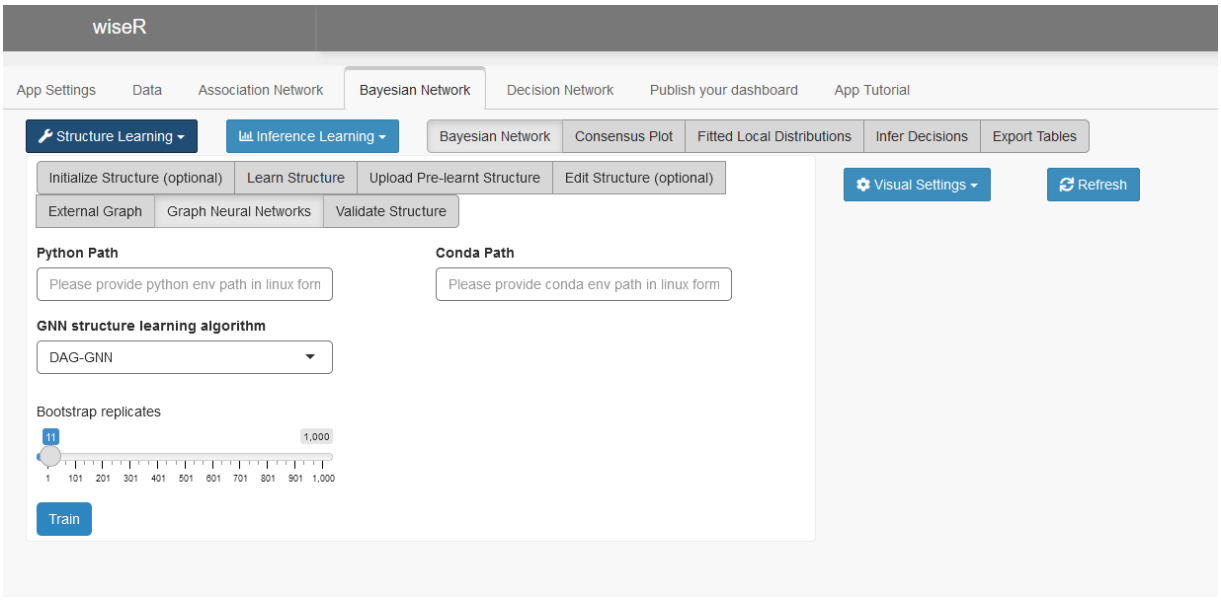


Figure S20. GNN based structure learning through the App.

3.7.1.7 Validate structure

This section enables the analyst to validate the learnt structure using 10-fold cross-validation or hold-out testing. The log-likelihood loss and network score outputs produced can then be used to judge the usability of the learned structure.

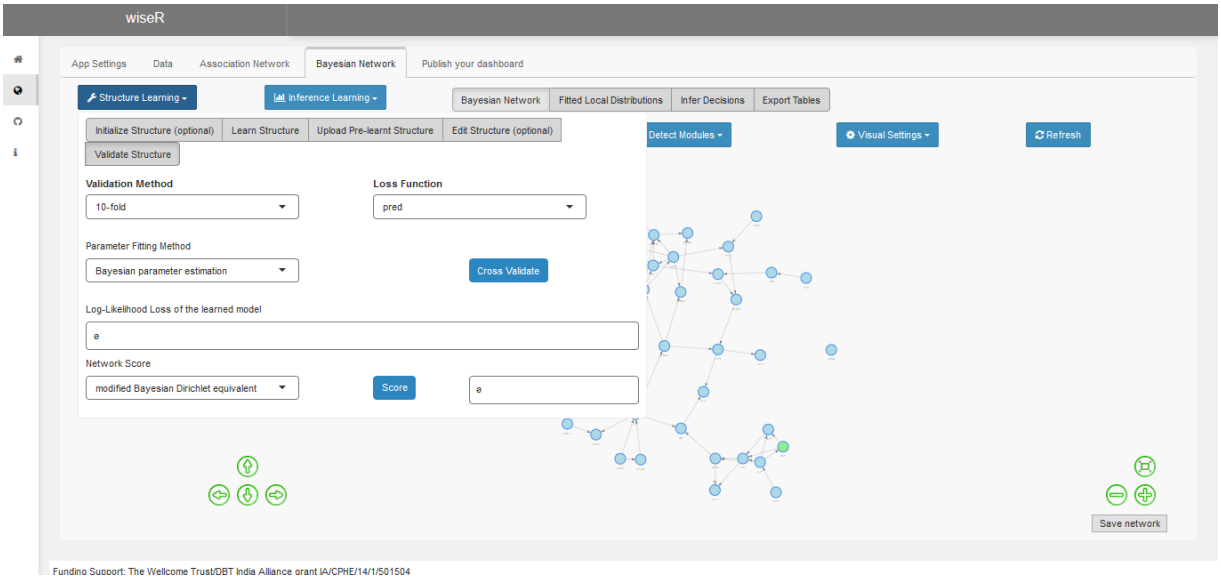


Figure S21. Validation of the learnt structure through cross-validation and hold-out testing.

3.7.2 Inference

After the parametrization of BN is complete, it can be queried to derive inferences and reason about the data. The inference learning menu enables the user to learn and explore conditional probability plots on the learned structure. The user can set an event and multiple pieces of evidence can be inserted as per the query. This crucial feature

enables the user to explore probability plots on event nodes in the network conditional on chain of evidence, which is the essence of decision making through Bayesian networks.

Inferences can be learnt through two methods using wiseR

Approximate Inference (Fast, Monte Carlo sampling of local structure). This option is useful for large datasets where exact inference is intractable. However, the sampling is implemented every time, hence adding incrementally to computational overhead over time. Also it yields slightly different results on each iteration. The app has the feature to re-compute approximate inferences for a specified number of times and to plot error bars in the inference.
Exact inference (Fast for small networks, slow for large networks, one-time computation)
Assuming that the User is working with large datasets, approximate inference is the default setting. This can be switched in the app. It is observed that the inference plots produced using approximate inference with error bars, is almost as accurate as exact inference, hence can enable robust reasoning and decisions on large datasets where exact inference is not possible.

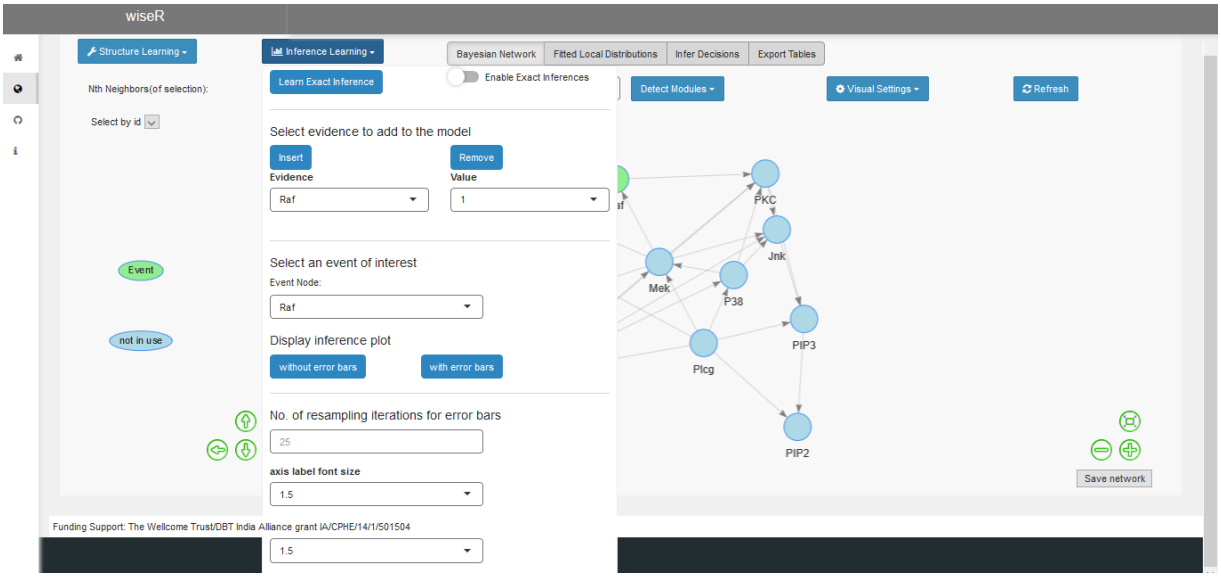


Figure S22. BN Inference. Approximate, Exact and Robustness.

Apart from structure and inference learning which are the workhorse of a BN analysis, the bayesian network tab allows options for visualization and exploration

- Bayesian network panel is used to explore the learned structure and is equipped with all the graph exploration features as available in association graph

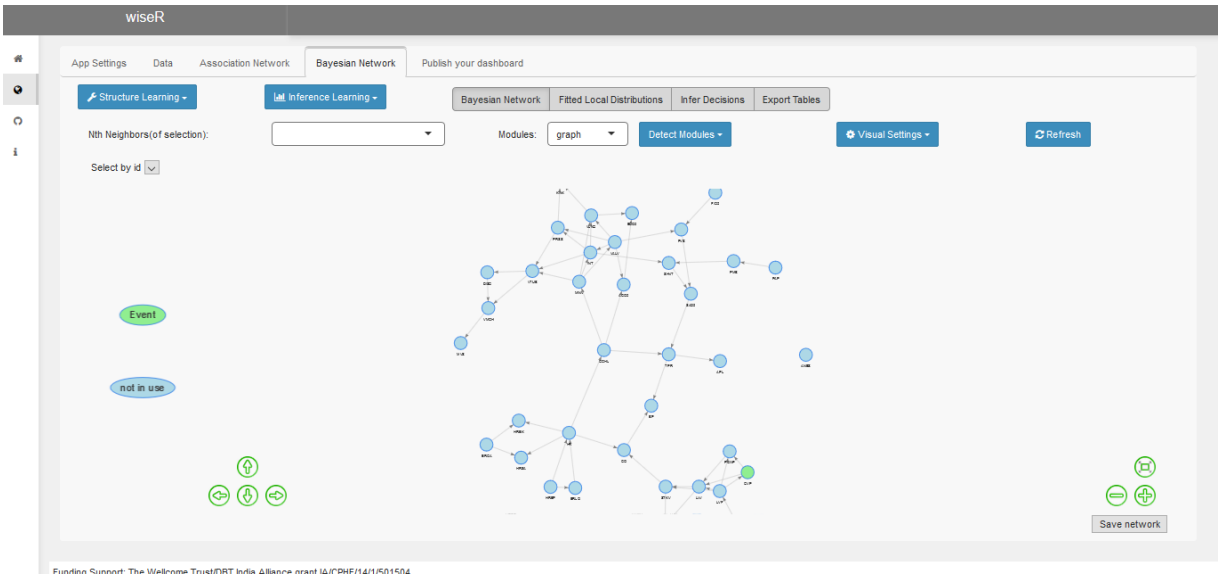


Figure S23. Example of a learnt bayesian graph

- Fitted local distribution panel is used to explore the local probability distribution tables of variables in the learned network structure

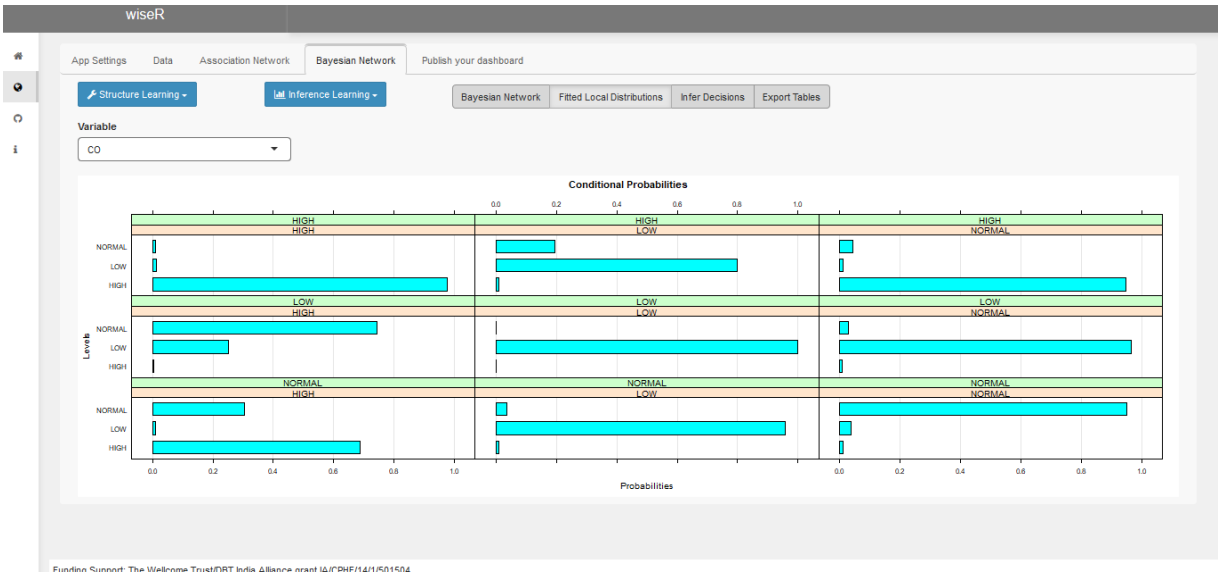


Figure S24. Local probability distribution tables

- Infer decisions are used to visualize the inference plots once the user has chosen the event and evidence nodes. It also has option to sort the plot axis and prune the no. of plot bars

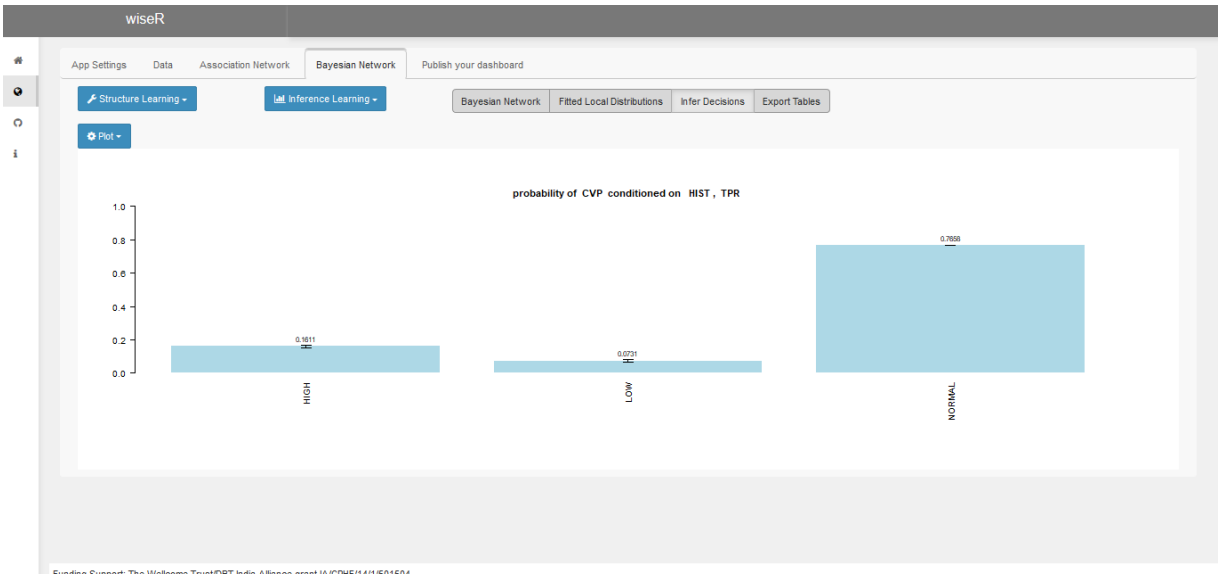


Figure S25. Inference plot

- Export tables section is used to visualize the network graph, blacklist-whitelist edges and variable wise validation results in tabular format. User can also download these tables as .CSV file

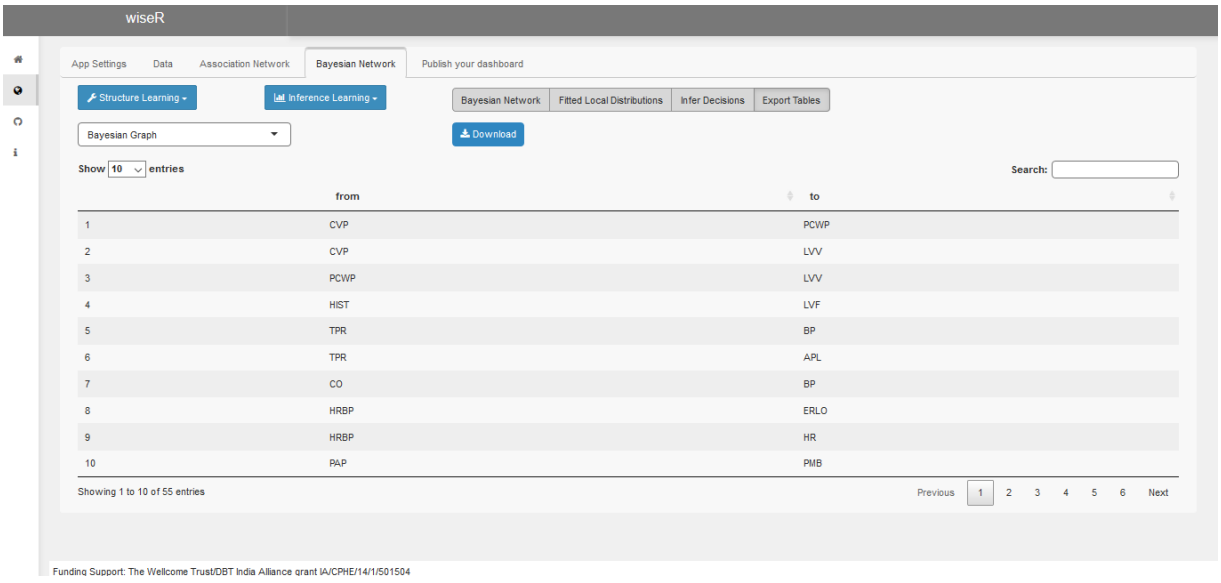


Figure S26. Visualization and tables available to be exported from the application.

3.8 Taking Decisions with the learnt network

Heads-up: This section relies upon JAGS installation on their system. JAGS can be downloaded from its official page.

As per our knowledge, there is no available open-source tool or a package in R that integrates structure learning with decision making. Influence diagrams have typically been defined by hand by experts. wiseR integrates these two capabilities to help the analyst arrive at the best (or next-best) policy for optimizing a payoff. For this purpose, the use can create a payoff node based on the variable that they wish to optimize (known as the utility node, e.g. Blood

Pressure optimization in the Intensive Care Unit). This brings up the table to enter the utility of each state in the range of -1 to 1. For example, a user may assign a negative utility to an abnormally low or high blood pressure event (-1) and a positive (+1) utility to normal blood pressure. Alternatively, if the user does not care as much about the high blood pressure as they would care about low blood pressure (e.g. shock in the ICU), they can assign a zero utility to the high blood pressure event. After this the user sets the decision nodes that are direct parents of the node to their utility node. If the policy from parents is expected to be obvious, they can go to grandparents or orders higher, however the effects may diminish. Finally the user runs the Monte-carlo simulation that displays the best policy along with the payoffs in a sorted table. This leap from making subjective inferences to policy advice makes wiseR a white-box approach to actionable, transparent and interpretable machine learning. A complete example is available towards the end of the tutorial for better understanding.

3.9 Publish your dashboard

Finally, the user can abstract the learning and decision making process into their own dashboard deployments. We consider this section to be a key highlight of wiseR. Publishing the analysis as an interactive dashboard helps reproducibility and effective communication. wiseR enables the user to produce a custom dashboard of their results as an R package.

The user has options to select the name and theme for the dashboard, and select the build option. wiseR builds the dashboard as a downloadable R/Shiny package which can be served as a web-application or launched as a local Shiny app. The custom dashboard is equipped with the most useful features of wiseR such as the interactive graph exploration and inference learning and visualization.

This feature saves authors the hassle of creating a dashboard for their findings, and helps propagate research output to the community in an efficient manner, thus helping bridging the gap between the research and actual use.

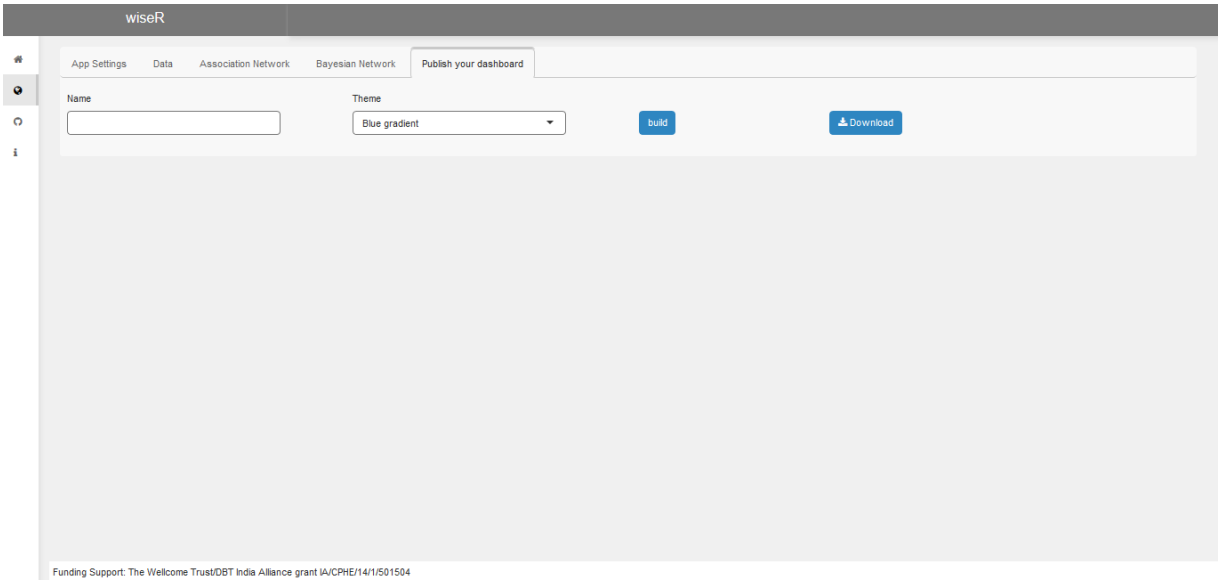


Figure S27. Publish your dashboard as a web-application or a local package

4 Comparison

- Open source (OS): Is the source code available ? If yes, what programming language?
- Target Audience (TA) :What is the target audience? Based on the source code availability and programming language used, the app can be targeted to end users, data scientists or both.
- Enabled Statistical/ML extensions (MLE) : Does the backend language support AI/ML libraries for customizations in those directions?
- Structure Learning (SL): Is structure learning possible?
- Bootstrapped Learning (BL): Is robust bootstrap structure learning possible?
- Interventional Data(INT): Can it handle interventional data and incorporate it in structure learning?
- Cross Validation(CV): Does the app provide means for cross validation and evaluation of learned models?
- Parameter learning(PL): is parameter learning possible using the app?
- Informed Layout(IL): is network visualization using a wide variety of efficient and targeted graph layouts for exploration ?
- Community Detection(CD) : are subgraphs based on communities/modules detected in the graph available for visualization and exploration?
- Inference with Confidence (IC): Can the app perform inference learning with confidence intervals ?
- Chained Inferences(CI): can the app handle multi-evidence inference learning?
- Exact Inference(EI): can the app perform exact inference learning?
- Approximate Inference with error bars(AIE): can the app perform much faster approximate inferences, verified over multiple iterations and produce final results with error bars(Much faster and equally accurate than exact inference on large datasets)?
- Decision Networks (DN): Does the app learn decision networks for policy optimization on specific utility nodes ?
- Graph Neural Network Structure (GNN'S): Does the graph have provision for GNN based structure Learning?
- Parallel processing(PP) : Does the app enable the user to run the process in parallel for faster runtime?
- Free: Is the app free for public use?

A comparison of wiseR and other structure learning and bayesian network applications

Name	OS	TA	MLE	SL	BL	INT	CV	PL	IL	CD	IC	CI	EI	AIE	DN	GNNL	PP	Free
wiseR	Yes (R)	Data Scientist,End user	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Bayesian Networks	Yes (R)	Data Scientist, End user	Y	Y	N	N	N	Y	N	N	N	N	N	N	N	N	N	Y
BayesiaLab	No	End User	N	Y	N	N	N	Y	N	N	N	Y	Y	N	N	N	N	N
Bayes Server	No	End User	N	Y	N	N	N	Y	N	N	N	Y	Y	N	N	N	N	N
BNJ	Yes (Java)	End User	N	N	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y
CIspace	Yes	End User	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	Y

	(Java)																
OpenMarkov	Yes (Java)	End User	N	Y	N	N	N	Y	N	N	N	Y	Y	N	N	N	N	Y
Sam Iam	No	End User	N	Y	N	N	N	N	N	N	N	N	Y	N	N	N	N	Y
UnBBayes	Yes (Java)	End User	N	Y	N	N	N	Y	N	N	N	Y	Y	N	N	N	N	Y

5 Example-1: Bayesian Inference

This section provides a hands-on example to conduct a step by step analysis to replicate one of the landmark results obtained using data-driven Bayesian Network learning and Inference, the Sachs dataset on yeast pathway modeling presented in Sachs et al. (2005), which utilizes bayesian learning on complex interventional data.

5.1 Steps

We upload the sachs interventional data which is in a space separated text file into the app. Variables are read as factors as it is a numerically encoded discrete data.

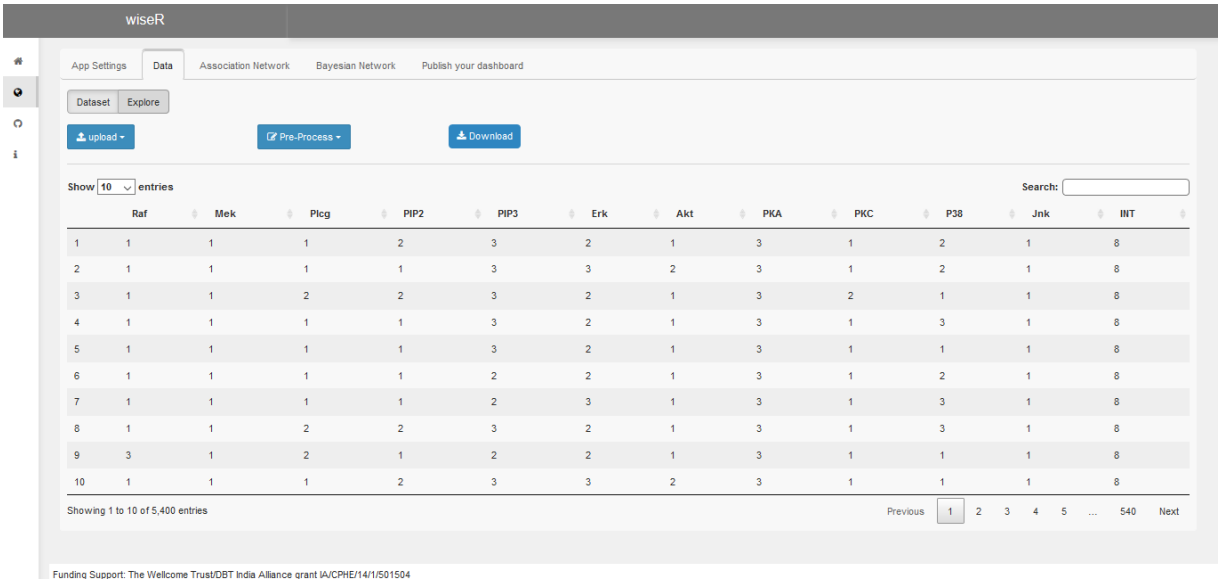


Figure S28. Sachs data uploaded

The variable INT holds interventional information, adjustment is made for the same in the preprocess menu. This will now be used as a part of Bayesian structure learning.

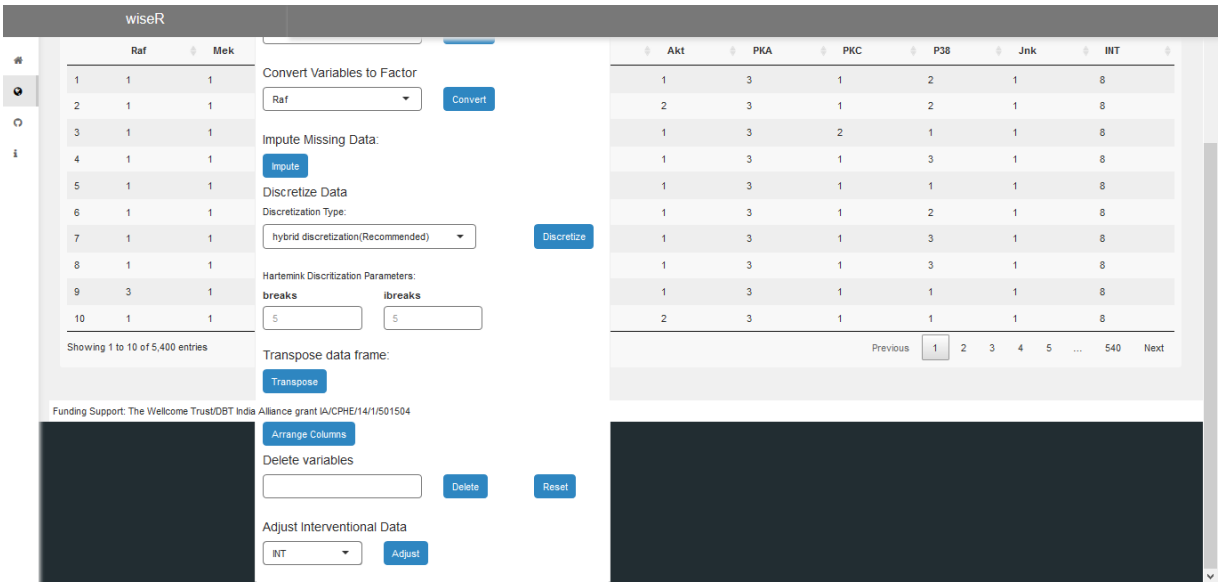


Figure S29. Handling interventional information

Next we learn a simple bayesian network on the data using Hill climbing algorithm and mbde network score. This graph will now be used to initialize structure learning using the tabu algorithm. This mechanism of graph initialization in structure learning is especially useful in score-based learning like Tabu, which prevents it from getting stuck in local maxima.

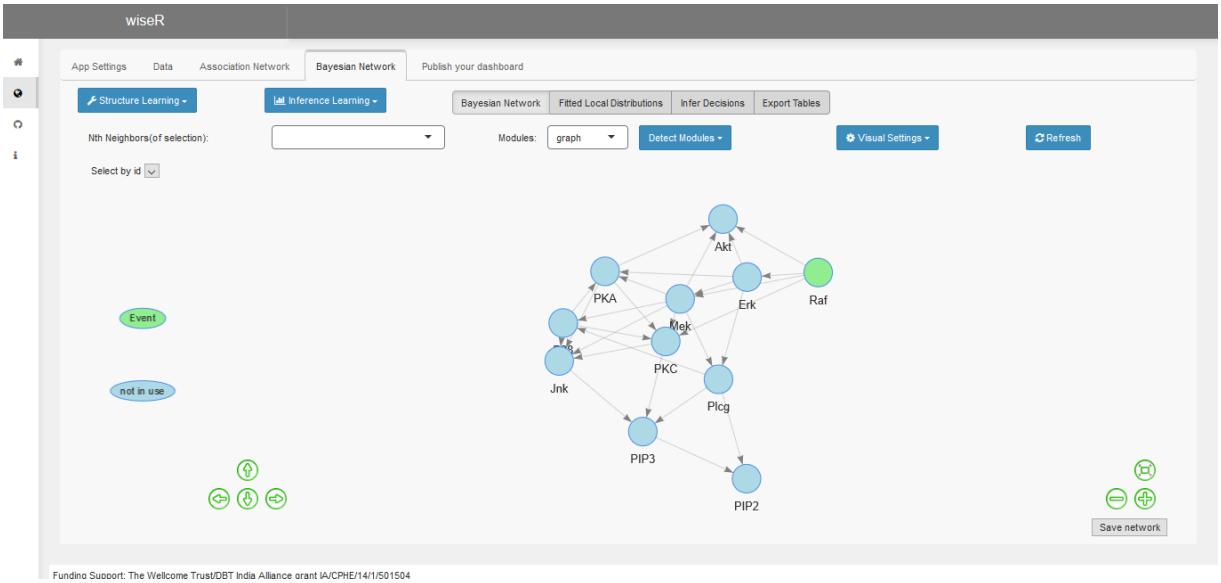


Figure S30. Learn structure using hill climbing

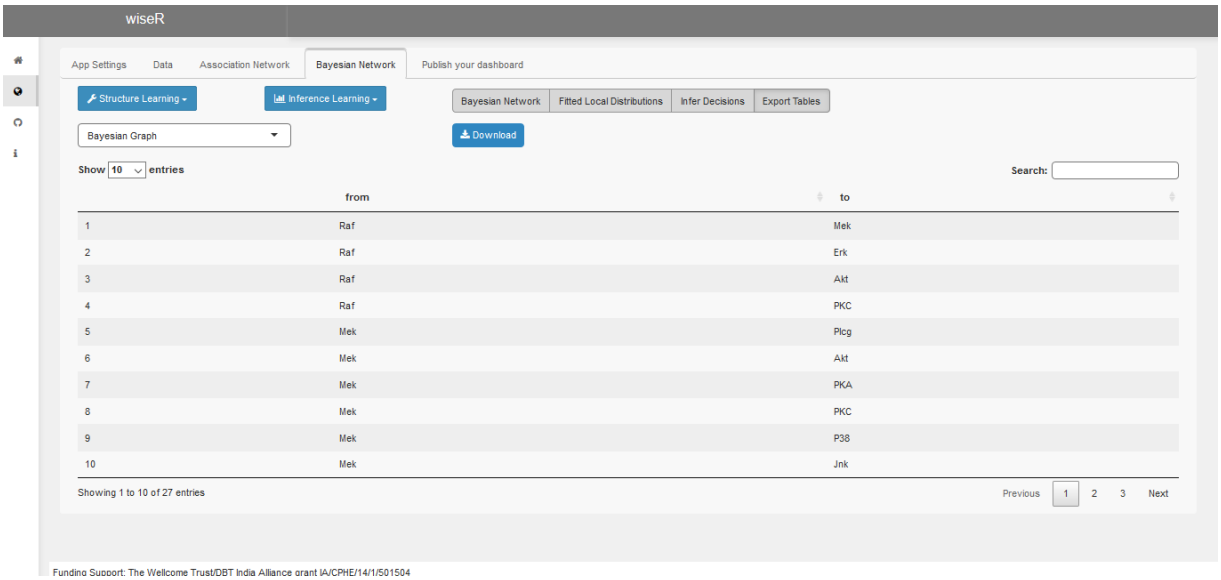


Figure S31. Download network graph as CSV file.

We now upload the graph as initialization for structure learning.

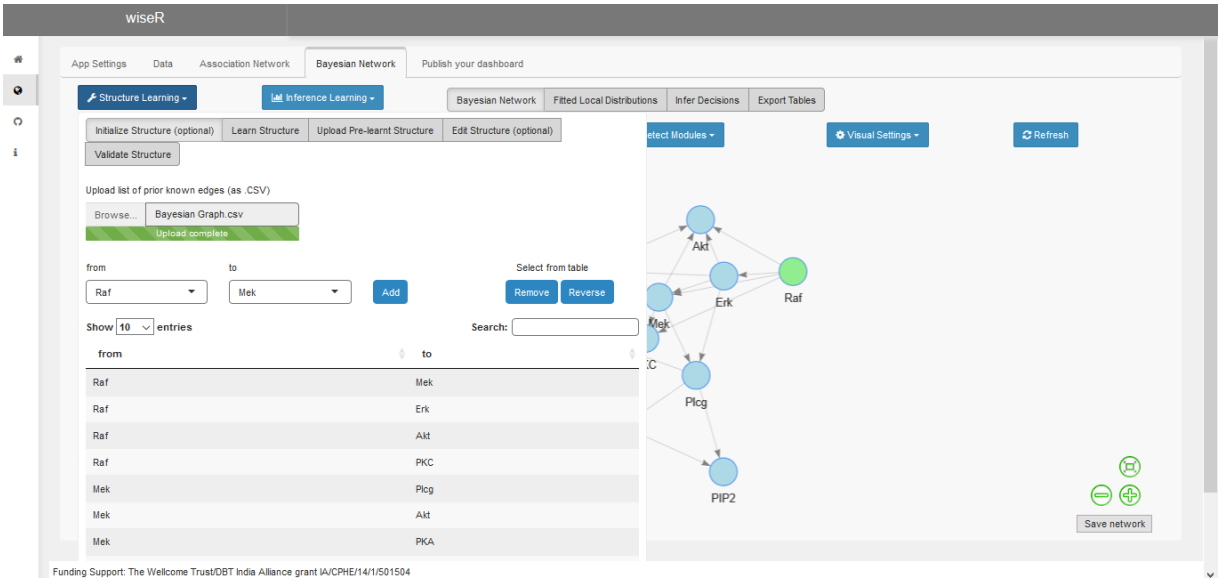


Figure S32. Initialize a new structure using the previously learnt graph

Next, we set the appropriate parameters for structure learning:

- Algorithm: tabu
- Network score: modified bayesian dirichlet equivalent
- ISS(imaginary sample size): 15
- Parameter fitting algorithm: bayesian parameter estimation

This performed a direct structure learning without bootstraps using tabu, to replicate the results as it was published in the study. However, we recommend that the user should select the bootstrap parameters and do a bootstrap learning, to produce more robust structures.

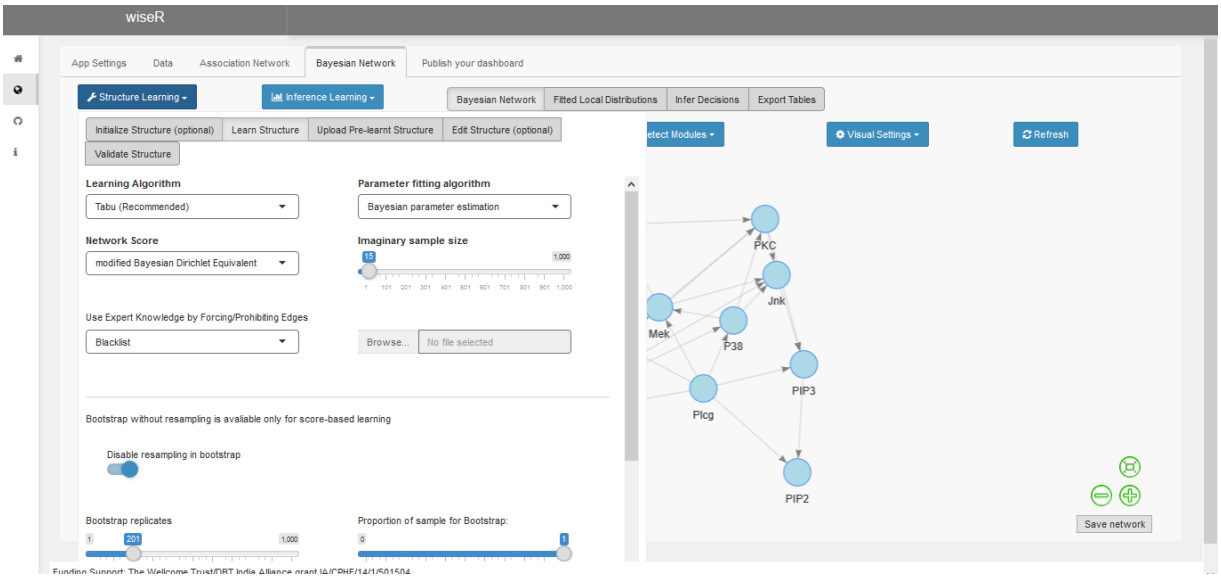


Figure S33. Structure learning using tabu algorithm on Sachs data

The final learnt structure replicates the results of the original paper and has all the validated arcs presented therein. It also discovered additional arcs which were not present in the original paper and may be useful for future research and exploration

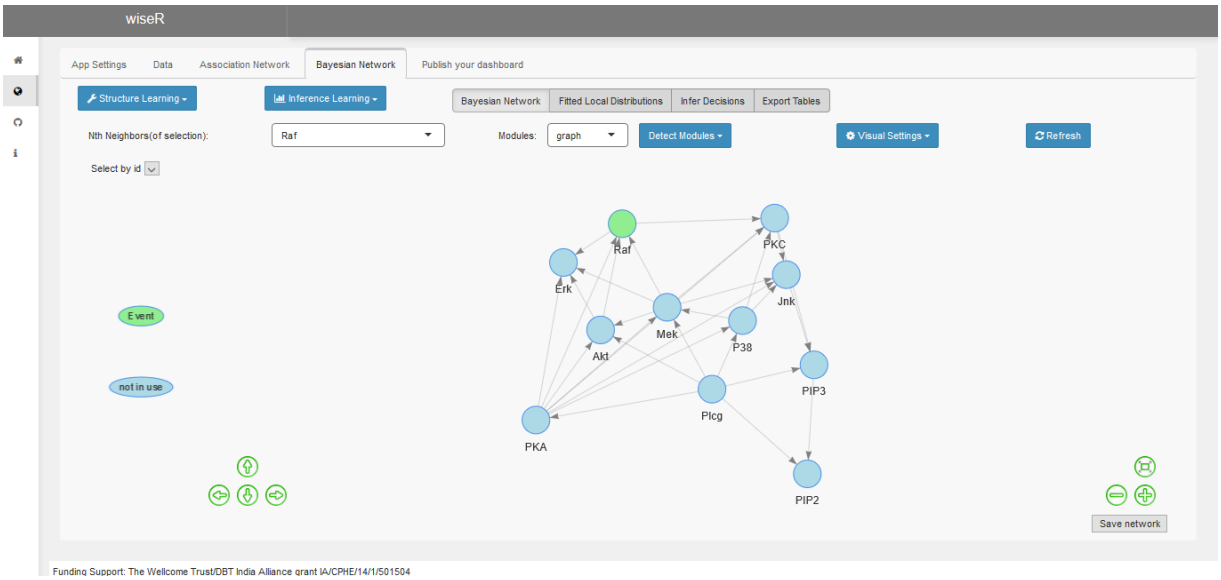


Figure S34. Final learned structure

We now confirm exact inference results with approximate inference with error bars to show that the unique approximate inference mechanism provided in the app produces as accurate results as exact inference without computational constraints in case of large structures. For this purpose we set the node ‘Akt’ as an event conditional

on ‘Erk = 1’. In the probability plots produced we can clearly see that approximate inference with error bars produces as accurate results as exact inference within 0.05 range of error, which is negligible

5.2 Exact Inference

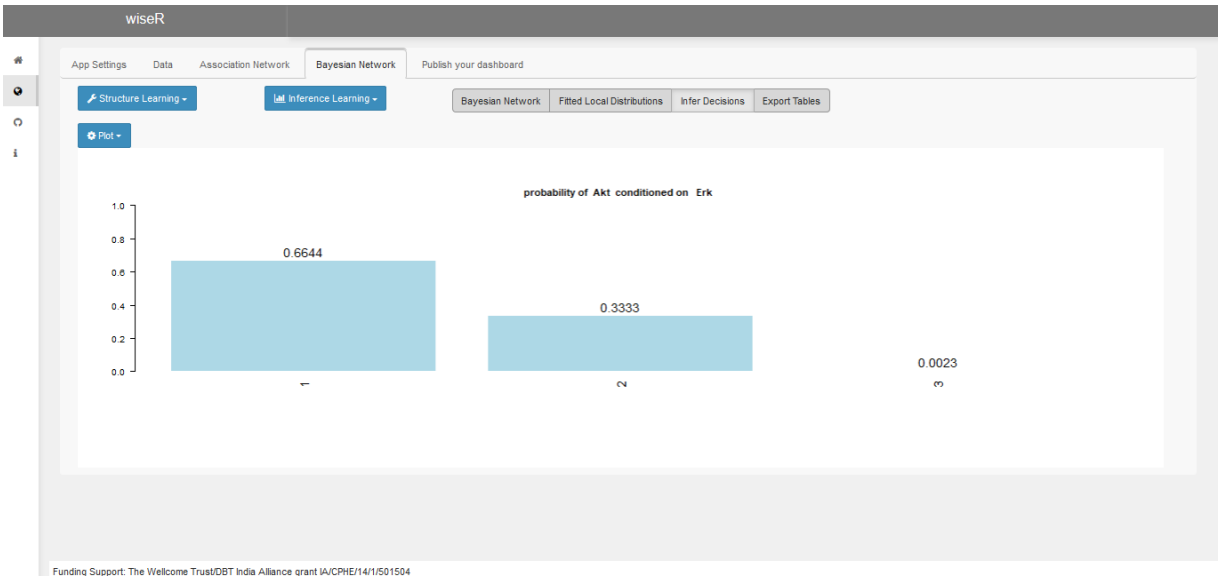


Figure S35. Exact inference plot

5.3 Approximate Inference with error bars

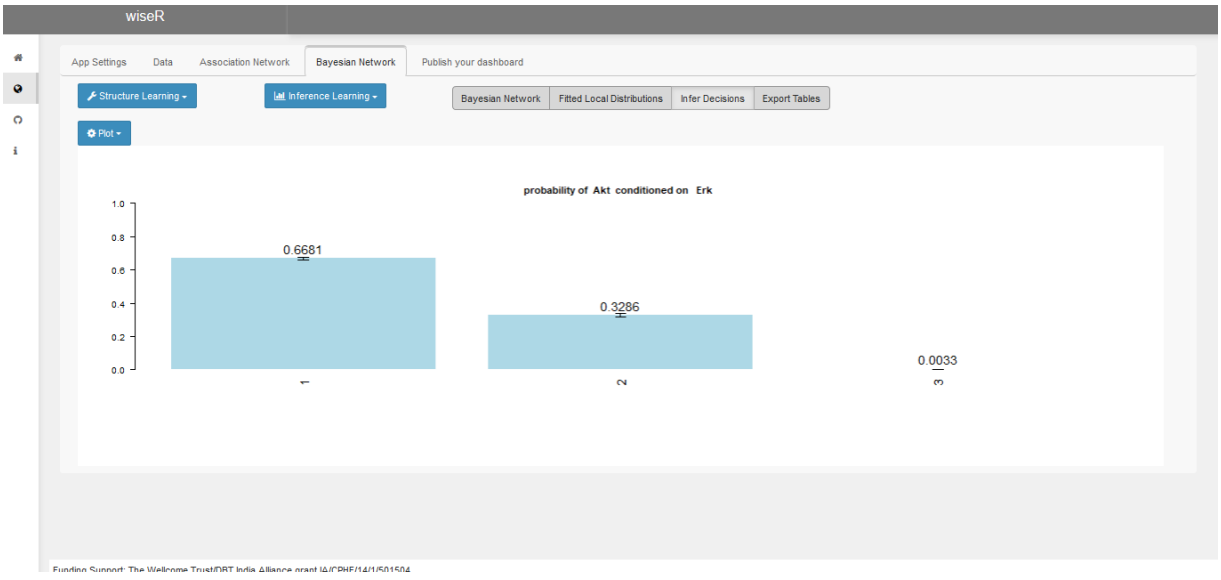


Figure S36. Approximate inference plot

This concludes the walk-through example for replicating a BN analysis using wiseR

6 Example-2: Decision Network

This section provides a hands-on example to learn decision networks and policy optimization through wiser. We use the Alarm dataset available in the pre-loaded datasets of the app.

We upload the already discretized and complete Alarm dataset from the app

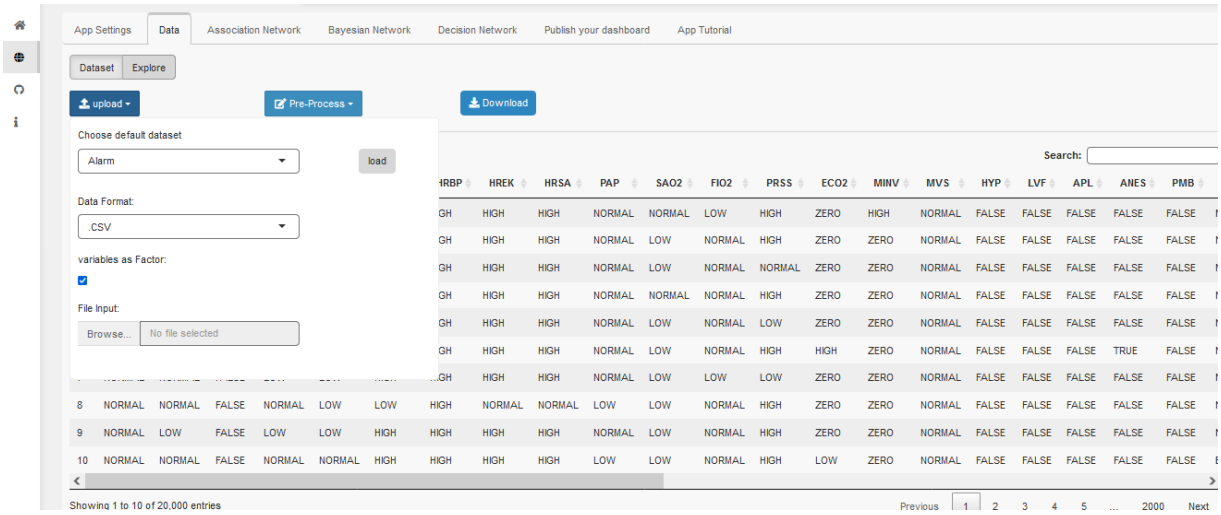


Figure S37. Alarm Data uploaded

We learn the Network structure using the following parameters:

- Algorithm: Hill Climbing
- Network Score: Akaike Information Criterion
- Parameter Fitting:Bayesian Parameter Estimation
- Blacklisting: No
- Bootstrap Model: Yes
- No. of Bootstrap:101
- Edge Strength: >0.51
- Direction Strength: >0.51
- Sample Portion: 1

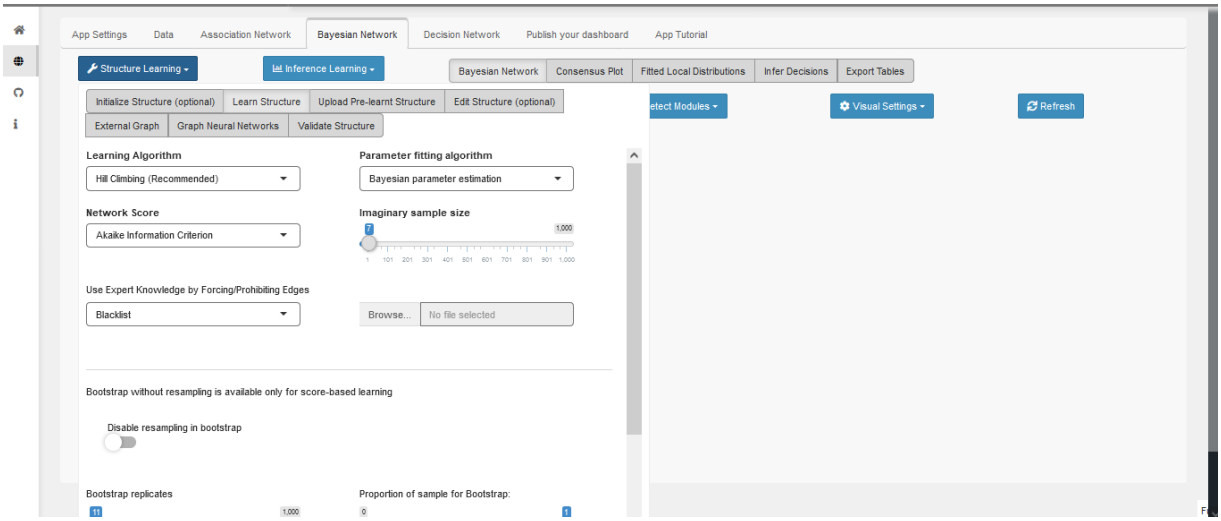


Figure S38. Structure Learning Parameters

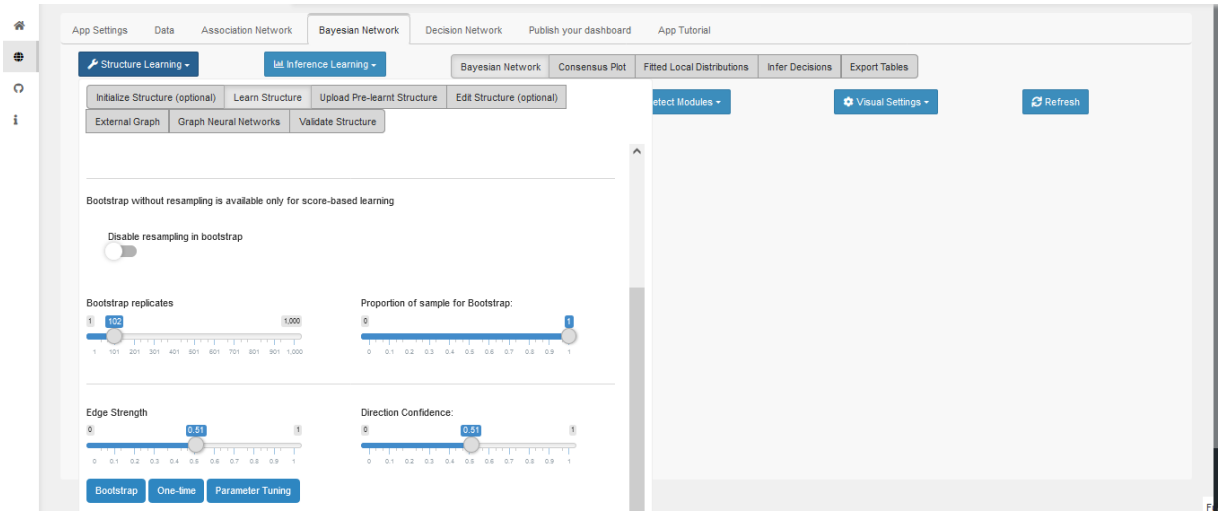


Figure S39. Bootstrap Parameters

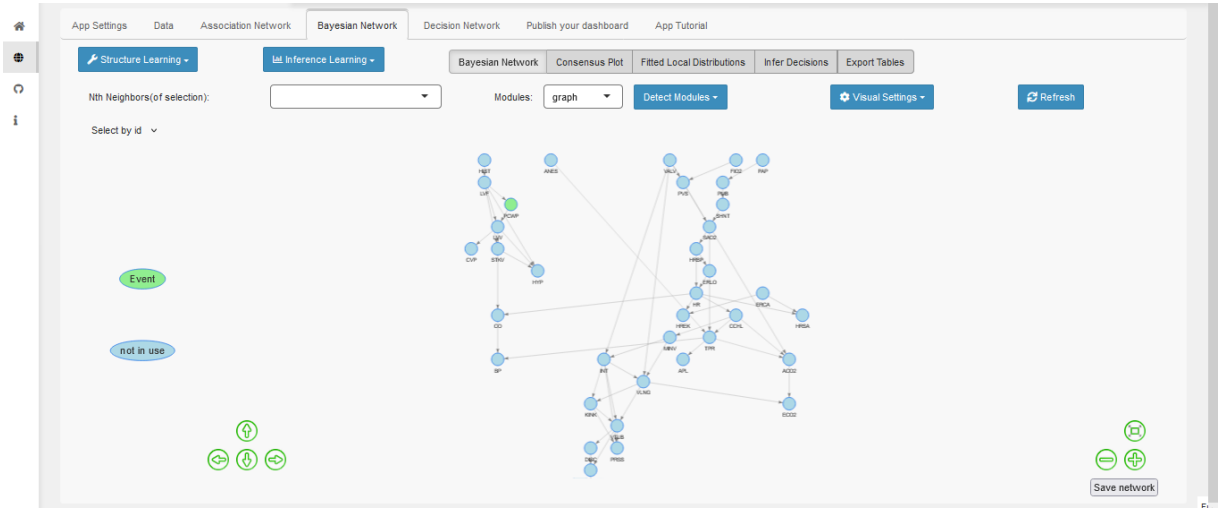


Figure S40. Learnt Bayesian Structure

As shown in the network the BP (blood pressure) has 2 causal parents the CO(cardiac output) and TPR (Total Peripheral Resistance)

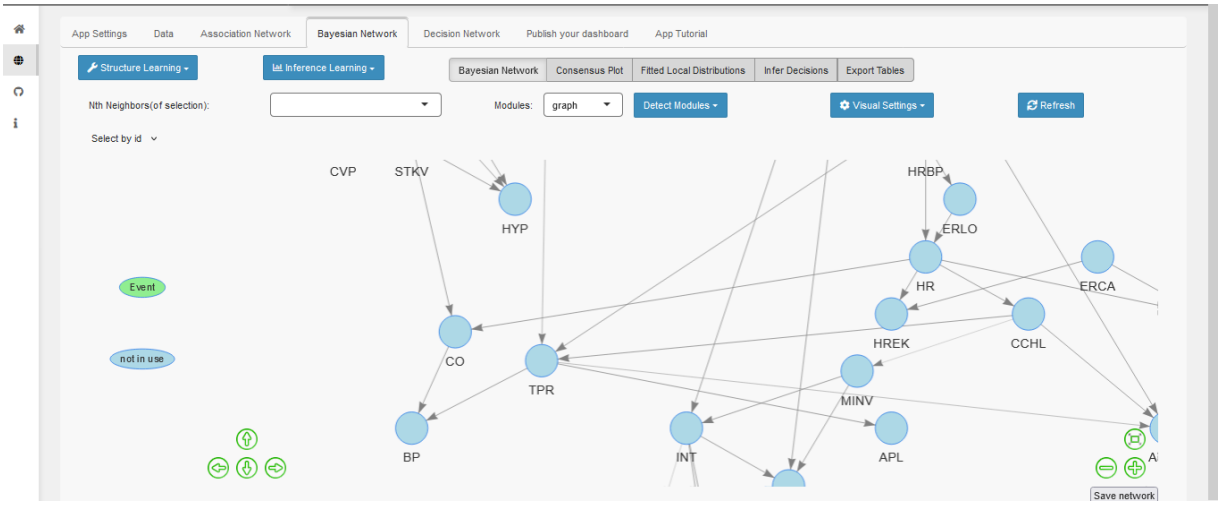


Figure S41. Nodes of interest in our structure

The aim of our decision network now is to find optimal values of CO and TPR to get a NORMAL value of BP. Thus we set our BP as utility node, payoffs of 1 for NORMAL, -0.5 for HIGH and -1 for LOW BP.

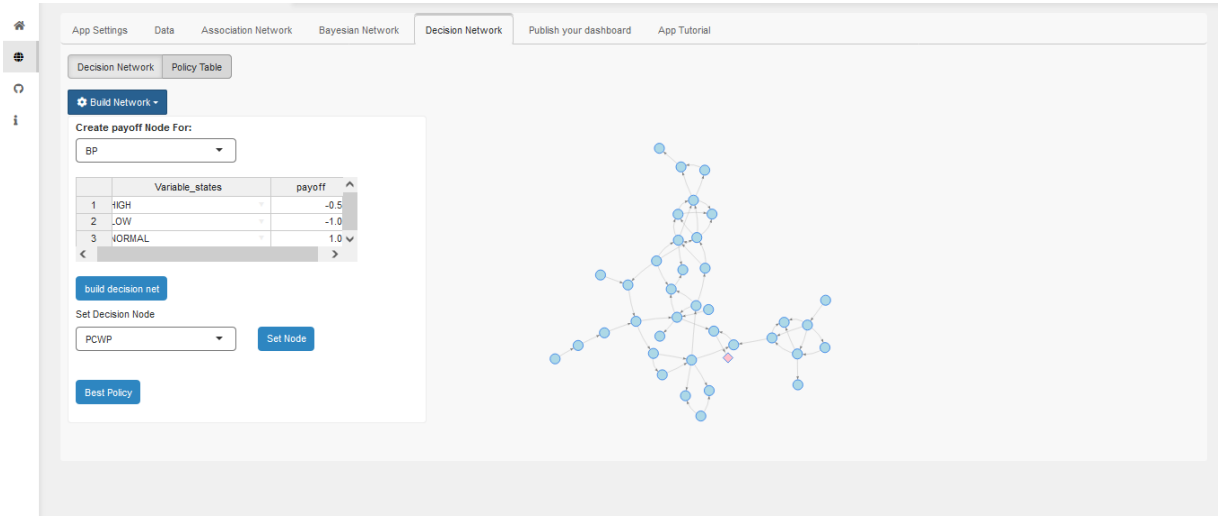


Figure S42. setting up utility node, corresponding payoffs and building the decision network

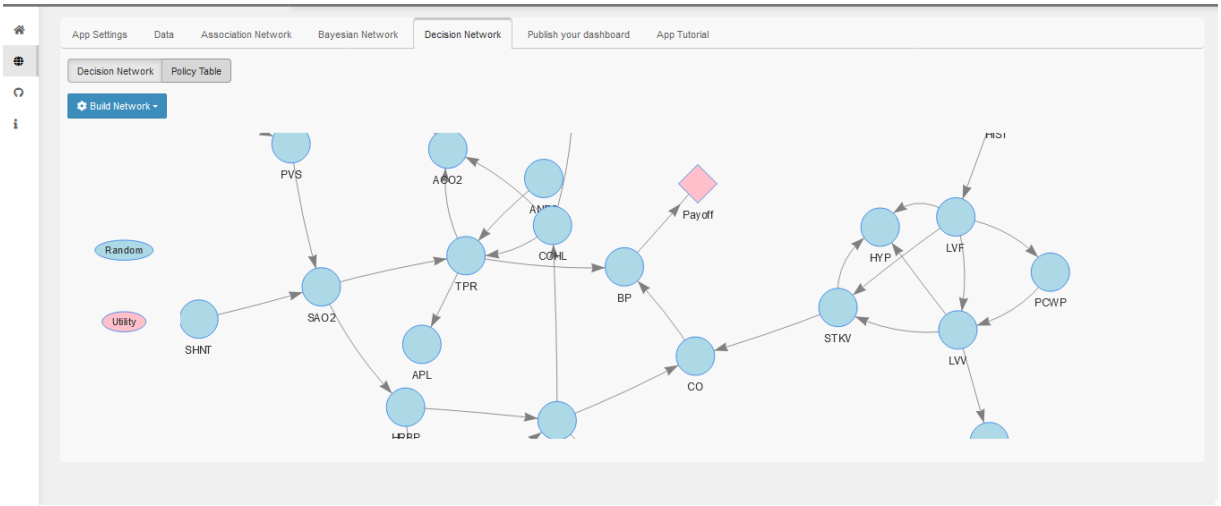


Figure S43. Payoff node in the Decision Network

Now we set CO and TPR as the decision nodes in the network. We aim to learn an optimal policy for both to maximize payoff on utility nodes.

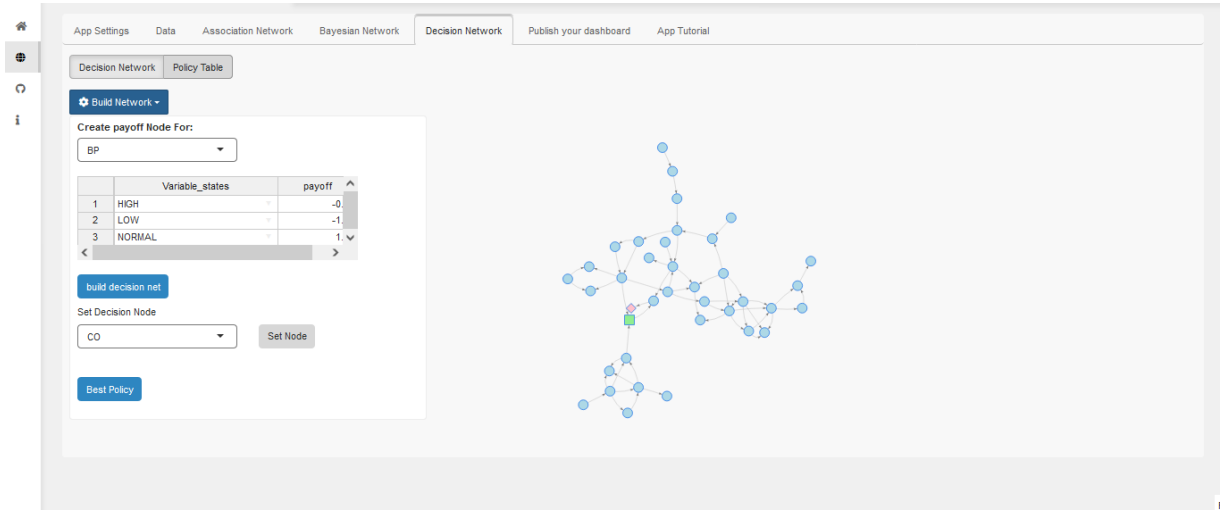


Figure S44. Setting Decision node for policy learning

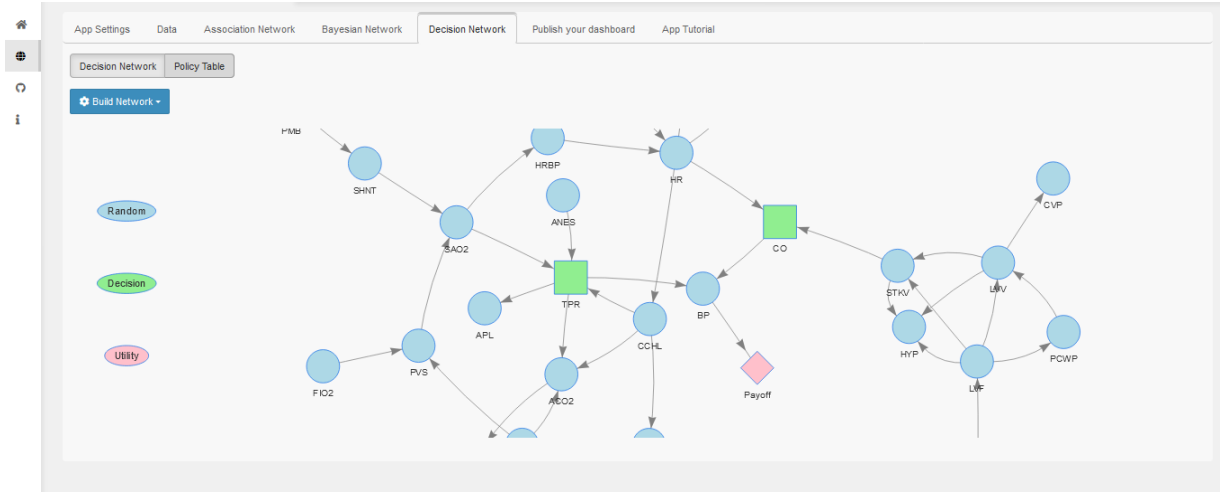


Figure S45. Final Layout of decision network depicting the utility and decision nodes

Finally from policy learning we can see that NORMAL values for both CO and TPR ensure maximum payoff and thus maximum probability of NORMAL BP. It is also interesting to note that a HIGH TPR more strongly affects high probability of NORMAL BP as compared to CO. Thus verifying the knowledge already available in the medical domain and validating the utility of wiser as a policy learning platform.

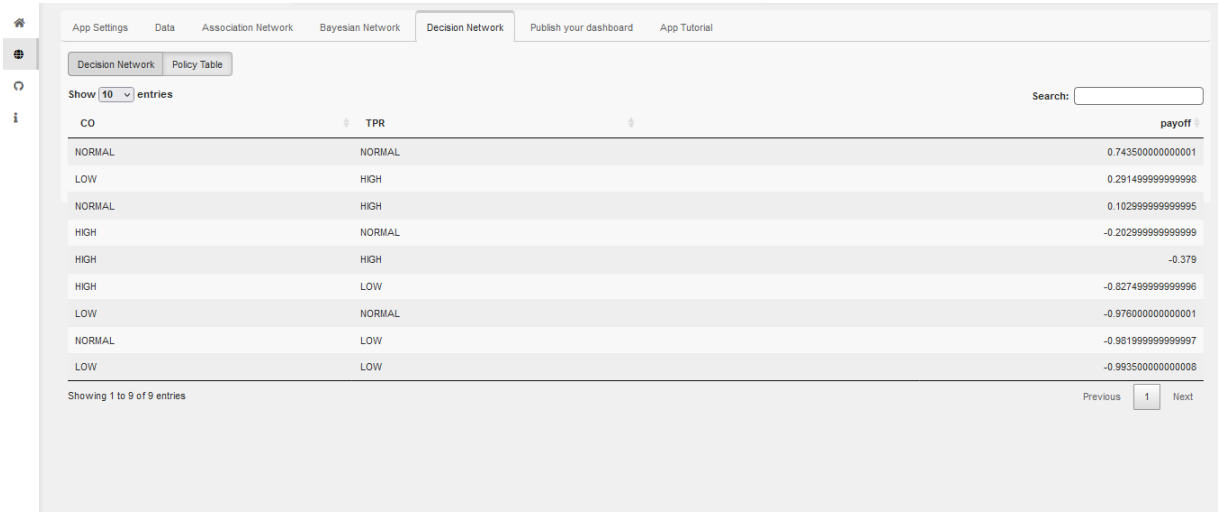


Figure S46. Learned policy table with corresponding payoffs

This concludes the walk-through example for replicating a Policy learning framework using Decision Networks using wiseR

7 wiseR Libraries

- RBGL (Vince Carey, Li Long and R. Gentleman (2017). RBGL: An interface to the BOOST graph library. R package version 1.52.0. http://www.bioconductor.org)
- graph (R. Gentleman, Elizabeth Whalen, W. Huber and S. Falcon (2017). graph: graph: A package to handle graph data structures. R package version 1.54.0.)
- bnlearn (Marco Scutari (2010). Learning Bayesian Networks with the bnlearn R Package. Journal of Statistical Software, 35(3), 1-22. URL http://www.jstatsoft.org/v35/i03/ .)
- rhandsontable (Jonathan Owen (2018). rhandsontable: Interface to the 'Handsontable.js' Library. R package version 0.3.6. https://CRAN.R-project.org/package=rhandsontable)
- shiny (Winston Chang, Joe Cheng, JJ Allaire, Yihui Xie and Jonathan McPherson (2017). shiny: Web Application Framework for R. R package version 1.0.5. https://CRAN.R-project.org/package=shiny)
- shinydashboard (Winston Chang and Barbara Borges Ribeiro (2017). shinydashboard: Create Dashboards with 'Shiny'. R package version 0.6.1. https://CRAN.R-project.org/package=shinydashboard)
- dplyr (Hadley Wickham, Romain Francois, Lionel Henry and Kirill Müller (2017). dplyr: A Grammar of Data Manipulation. R package version 0.7.4. https://CRAN.R-project.org/package=dplyr)
- visNetwork (Almende B.V., Benoit Thieurmel and Titouan Robert (2018). visNetwork: Network Visualization using 'vis.js' Library. R package version 2.0.3. https://CRAN.R-project.org/package=visNetwork)
- shinyWidgets (Victor Perrier and Fanny Meyer (2018). shinyWidgets: Custom Inputs Widgets for Shiny. R package version 0.4.1. https://CRAN.R-project.org/package=shinyWidgets)
- missRanger (Michael Mayer (2018). missRanger: Fast Imputation of Missing Values. R package version 1.0.2. https://CRAN.R-project.org/package=missRanger)
- tools (R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/ .)
- shinyalert (Dean Attali and Tristan Edwards (2018). shinyalert: Easily Create Pretty Popup Messages (Modals) in 'Shiny'. R package version 1.0. https://CRAN.R-project.org/package=shinyalert)
- shinycssloaders (Andras Sali (2017). shinycssloaders: Add CSS Loading Animations to 'shiny' Outputs. R package version 0.2.0. https://CRAN.R-project.org/package=shinycssloaders)

- rintrojs (Carl Ganz (2016). rintrojs: A Wrapper for the Intro.js Library. Journal of Open Source Software, 1(6), October 2016. URL http://dx.doi.org/10.21105/joss.00063)
- arules (Michael Hahsler, Christian Buchta, Bettina Gruen and Kurt Hornik (2018). arules: Mining Association Rules and Frequent Itemsets. R package version 1.6-1. https://CRAN.R-project.org/package=arules)
- psych (Revelle, W. (2018) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version = 1.8.4.)
- DescTools (Andri Signorell et mult. al. (2018). DescTools: Tools for descriptive statistics. R package version 0.99.24.)
- DT (Yihui Xie (NA). DT: A Wrapper of the JavaScript Library 'DataTables'. R package version 0.4.11. https://rstudio.github.io/DT)
- linkcomm (Kalinka, A.T. and Tomancak, P. (2011). linkcomm: an R package for the generation, visualization, and analysis of link communities in networks of arbitrary size and type. Bioinformatics 27 (14), 2011-2012.)
- igraph (Csardi G, Nepusz T: The igraph software package for complex network research, InterJournal, Complex Systems 1695. 2006. http://igraph.org)
- parallel (R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.)
- snow (Luke Tierney, A. J. Rossini, Na Li and H. Sevcikova (2016). snow: Simple Network of Workstations. R package version 0.4-2. https://CRAN.R-project.org/package=snow)
- shinyBS (Eric Bailey (2015). shinyBS: Twitter Bootstrap Components for Shiny. R package version 0.61. https://CRAN.R-project.org/package=shinyBS)
- gRbase (Claus Dethlefsen, Søren Højsgaard (2005). A Common Platform for Graphical Models in R: The gRbase Package. Journal of Statistical Software, 14(17), 1-12. URL http://www.jstatsoft.org/v14/i17/.)
- gRain (Søren Højsgaard (2012). Graphical Independence Networks with the gRain Package for R. Journal of Statistical Software, 46(10), 1-26. URL http://www.jstatsoft.org/v46/i10/.)
- RCy3 (Ono K, Muetze T, Kolishovski G, Shannon P, Demchak, B. CyREST: Turbocharging Cytoscape Access for External Tools via a RESTful API [version 1; referees: 2 approved]. F1000Research 2015, 4:478.)
- BiocManager (Martin Morgan (2018). BiocManager: Access the Bioconductor Project Package Repository. R package version 1.30.1. https://CRAN.R-project.org/package=BiocManager)
- HydeNet (Jarrod E. Dalton and Benjamin Nutter (2018). HydeNet: Hybrid Bayesian Networks Using R and JAGS. R package version 0.10.7. https://CRAN.R-project.org/package=HydeNet)

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

Wrong Side Vehicle Detection

Anirudh Goyal
Department of Electrical Engg.
Delhi Technological University
New Delhi, India
anirudhgoyal136@gmail.com

Antriksh Mangal
Department of Electrical Engg.
Delhi Technological University
New Delhi, India
antrikshmangal@gmail.com

A.R.Kulkarni
Department of Electrical Engg.
Delhi Technological University
New Delhi, India
ashishkulkarni@dtu.ac.in

2021 International Conference on Intelligent Technologies (CONIT) | 978-1-7281-8583-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONIT51480.2021.9498416

***Abstract*—With the advent of Intelligent Transportation Systems, traffic management can be made more robust and advanced. This can alleviate problems such as wrong side driving, which is one of the leading causes of traffic congestion and road accidents, thus improving the road transport efficiency. In this paper, we present a novel solution to reduce the number of accidents caused by wrong side vehicles by alerting the other road users and warning the driver to revert back. Present day solutions are expensive and difficult to setup. Using low-cost hardware and pragmatic design, we are able to provide economic viability to our solution.**

***Index Terms*—ITS, ANPR, OCR, Tesseract, CNN, Arduino UNO, Raspberry Pi, IR Proximity Sensor, DMD Display**

I. Introduction

Wrong side driving is defined as the movement of motor vehicles against the flow of traffic often resorted to for reducing commute time. We carried out a field survey near the road-intersection at the entry gate of a school in New Delhi. We observed rampant traffic rule violations of kinds such as red light jumping, over-speeding, halting on/after zebra-crossing and wrong side driving, which was of the most dangerous nature. It is one of the leading causes of road accidents [1]. According to a recent report by Ministry of Road Transport Highways, Government of India in the year 2018, driving on wrong side and lane indiscipline resulted in 24,781 road accidents due to which 8,764 people were killed and 24,100 were injured. These increased to 24,431 road accidents leading to 9,201 deaths and 24,628 severe injuries in the year 2019. The report also mentions that wrong side driving is the second most important cause accounting for 5.4 percent of the road accidents, 6.1 percent of total deaths and 5.5 percent of total injuries.

We propose a wrong side vehicle detection system wherein IR proximity sensors are installed on the side of the road. Upon detecting the vehicle movement in the wrong direction, our ANPR system is triggered to display the vehicle registration number on the LED panel simultaneously alerting other road users.

II. Hardware Description

Our system uses low cost hardware components comprising of IR proximity sensors, camera, processing unit and DMD display panel.

A. IR Proximty Sensor

The IR Proximity Sensor is a multipurpose infrared sensor which has been used for vehicle detection. Commercially available sensors have a maximum detection range of 85 cm.

The sensor outputs a logic one (+24V DC) when an object is placed in front of the sensor and a logic zero (0V), when there is no object in front of the sensor. In order to interface with Arduino, buck converters have been used to bring down the voltage level to +5V DC. A 24V 5A Switched Mode Power Supply(SMPS) provides power to the sensors.

B. Camera

The network camera used in this system has a CMOS image sensor with a lens of F1.0 aperture. It provides a resolution of 1920x1080 pixels at a frame rate of 30fps. The stream can be accessed over TCP/IP, HTTP, HTTPS, RTP, and RTSP protocols. The camera draws a current of 0.5 A at 12V DC thereby consuming a maximum power of 6 W.

C. Processing Unit

The processing unit comprises of an Arduino UNO Module in serial communication with a Raspberry Pi 4 single-board computer.

Arduino UNO is an ATmega328P microchip based microcontroller equipped with 14 digital I/O pins (six capable of PWM output), 6 analog I/O pins, and is programmable with the Arduino IDE. It is powered by a USB cable connected serially to the Raspberry Pi consuming a maximum current of 300mA while operating at 5V.

Raspberry Pi 4 Model B is a small single board computer based on a 1.5 GHz 64-bit quad core ARM Cortex-A72 processor. The VideoCore VI on-board GPU allows state-of-the-art object detection model architectures to be run for executing the ANPR algorithm[2]. The Pi 4 is powered via a USB-C port from the official Raspberry Pi adapter.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:59:34 UTC from IEEE Xplore. Restrictions apply.

D. DMD Display

P10 LED Dot-Matrix Display(DMD) Modules are used to display the vehicle license plate number as they are the most suitable for outdoor applications, and can be easily viewed from a distance of 12 meters. A single module consists of 512 high brightness LEDs mounted on a plastic housing. The size of the module is 32*16 which means that there are 32 LEDs in each row and 16 LEDs in each column.

To display the number which ranges from 9-11 characters (according to Indian RTO regulations), we have connected 3 DMD modules in a row. Each module has a brightness of 3500-4500 nits consuming a maximum power of 20W at 5V DC with an IP65 waterproof rating. The module is controlled from the digital I/O pins on Arduino board using a 16-pin flat ribbon connector.

The processing unit along with the power supply is placed in a plastic case. The entire setup is mounted on a cast iron frame of height 190 cms and weighs around 15kg.

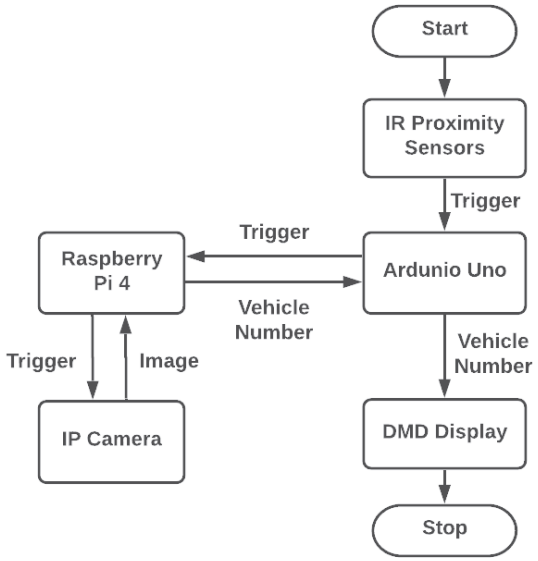


Fig.1 Complete Flow control

III. Design Description

A. Determining direction of motion

Two IR proximity sensors are placed 1 meter apart on the side of the road. When the vehicle approaches from either direction, the sequence of triggering of sensors(S1 and S2) determines the direction of vehicle movement [3]. We encountered two possible cases

Case 1: S1 outputs logic 1 followed by S2 and consequently S1 outputs logic 0 followed by S2. The algorithm interprets that the vehicle is moving in the direction opposite to the flow of traffic, i.e., in the wrong way.

Case 2: S2 outputs logic 1 followed by S1 and thereafter S2 outputs logic 0 followed by S1. The algorithm interprets that the vehicle is moving in the direction of flow of traffic.

When case 1 is detected, i.e., the vehicle is moving in the wrong direction, the Arduino sends a logic 1 to the Raspberry Pi GPIO pin 24 thus triggering the ANPR system. Additionally, the time difference between triggering of the two sensors is stored in a variable which can be used to calculate the speed of the vehicle since the distance between S1 and S2 is predefined.

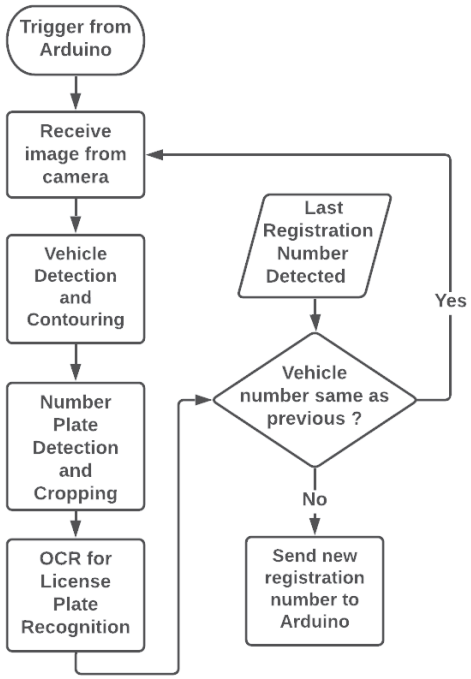


Fig.2 Raspberry Pi Process Flow Diagram

B. Automatic Number Plate Recognition

After receiving the trigger from Arduino, the Raspberry Pi initiates the ANPR algorithm [4], [5] on an image captured by the camera. The vehicle, which can be a car or a motorbike, is identified in the image using MobileNet [6] Single Shot Detector [7] (MobileNet SSD). MobileNet SSD is based on depth-wise separable convolutions that essentially splits the standard convolution into two layers, one layer for filtering and the other for combining. This drastically reduces the computation and model size.

YOLO architecture [8] is another object detection algorithm based on region proposal networks such as fast R-CNN which passes the image once through a fully convolutional neural network, splitting the input image into a grid and for each grid generating 2 bounding boxes and class probabilities for those bounding boxes [9].

However, for our purpose of vehicle detection being run on the Raspberry Pi computer, MobileNet SSD having an input size of 300X300 produces a higher mean average precision (mAP) metric with a generous frame rate, allowing close to real-time frame rate. After creating a bounding-

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:59:34 UTC from IEEE Xplore. Restrictions apply.

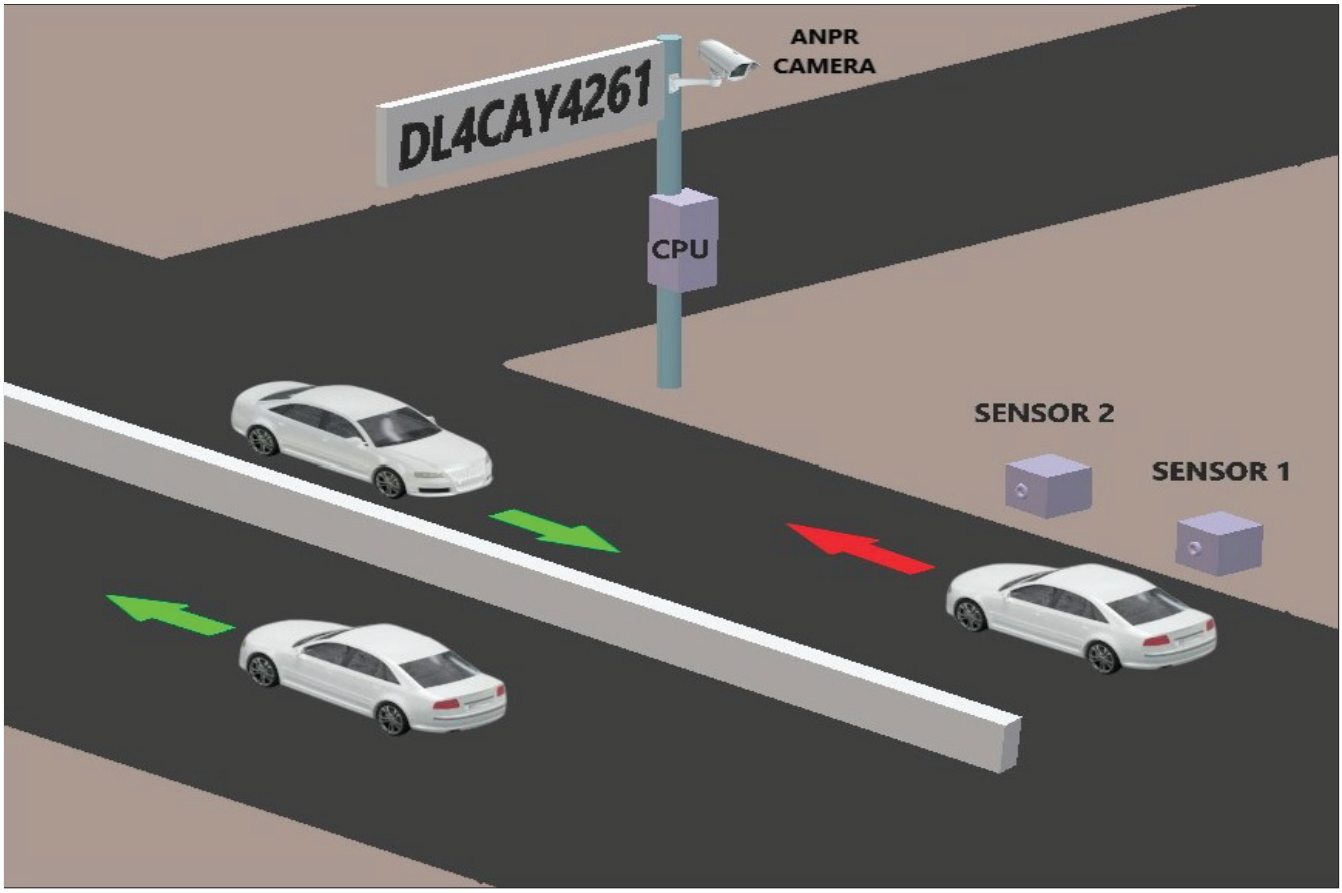


Fig.3 WSVD System Design. Green arrows indicate the correct direction of motion and red arrow indicates the wrong side moving vehicle. The vehicle registration number is read and displayed on the DMD module.

TABLE I
COMPARING DETECTION FRAMEWORKS

Detection Framework	Train	mAP	FPS
Fast R-CNN	2007+2012	70.0	0.5
Faster R-CNN VGG-16	2007+2012	73.2	.7
Faster R-CNN ResNet	2007+2012	76.4	5
YOLO	2007+2012	63.4	45
Mobilenet SSD300	**2007+2012**	**74.3**	**46**
Mobilenet SSD500	2007+2012	76.8	19

box around the vehicle in the image, it is cropped out and another MoblilenetSSD based object detection model fine-tuned for number plates is applied. The model outputs the number plate coordinates which are in turn used for the OCR algorithm. For choosing an appropriate Optical character recognition model, we draw a comparison between two OCR algorithms . Tesseract[10] is an optical character recognition software, released under the Apache License, now maintained by Google. It performs poorly on input images that are not preprocessed. Images need to be scaled up to at least 20 pixels and any rotation or skew must be rectified. Tesseract's binarization stage destroys low-frequency changes in the brightness, and dark borders are misinterpreted as characters.

The OCR algorithm used in Rayson Laroca et al. [11] comprises of character segmentation based on the YOLO-VOC network proposed by Montazzolli and Jung (CR-NET) [12] which is trained using License Plate patch (with a margin) and the characters coordinates as inputs. For character recognition, two CNNs have been used. For digit recognition, first four layers of the character segmentation CNN have been removed. However, for letter recognition the entire architecture of the character segmentation CNN has been used.

C. Displaying the Number

To display the vehicle license plate number, Dot-Matrix LED display module is used. The display is connected to Arduino digital I/O pins using 16-pin FRC connector and is powered separately using 5V DC power supply unit.

To interact with the display, DMD.h library is used along with TimerOne.h as it provides a high speed display connection straight to the pins and the SPI port. It includes a complete 5x7 pixel and 6x16 pixel font set, character routines for display and test pattern generation functions. Arduino receives the number generated by the OCR algorithm from Raspberry Pi using serial data communication

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:59:34 UTC from IEEE Xplore. Restrictions apply.

TABLE II
YOLO-VOC BASED CHARACTER SEGMENTATION CNN AS PROPOSED IN [11]

Layer	Filters	Size	Input	Output
1 conv	32	3 × 3/1	240×80×3	240×80×32
2 max		2 × 2/2	240×80×32	120×40×32
3 conv	64	3 × 3/1	120×40×32	120×40×64
4 max		2 × 2/2	120×40×64	60×20×64
5 conv	128	3 × 3/1	60×20×64	60×20×128
6 conv	64	1 × 1/1	60×20×128	60×20×64
7 conv	128	3 × 3/1	60×20×64	60×20×128
8 max		2 × 2/2	60×20×128	30×10×128
9 conv	256	3 × 3/1	30×10×128	30×10×256
10 conv	128	1 × 1/1	30×10×256	30×10×128
11 conv	256	3 × 3/1	30×10×128	30×10×256
12 conv	512	3 × 3/1	30×10×256	30×10×512
13 conv	256	1 × 1/1	30×10×512	30×10×256
14 conv	512	3 × 3/1	30×10×256	30×10×512
15 conv	30	1 × 1/1	30×10×512	30×10×30
16 detection				

and then displays it on the DMD module for a default time of 5 seconds so that the traffic-rule violator coming from the wrong direction can be warned. If another wrong side vehicle follows within 5 seconds then the previous number is overwritten. Simultaneously, the number is saved in local database along with the corresponding timestamp.

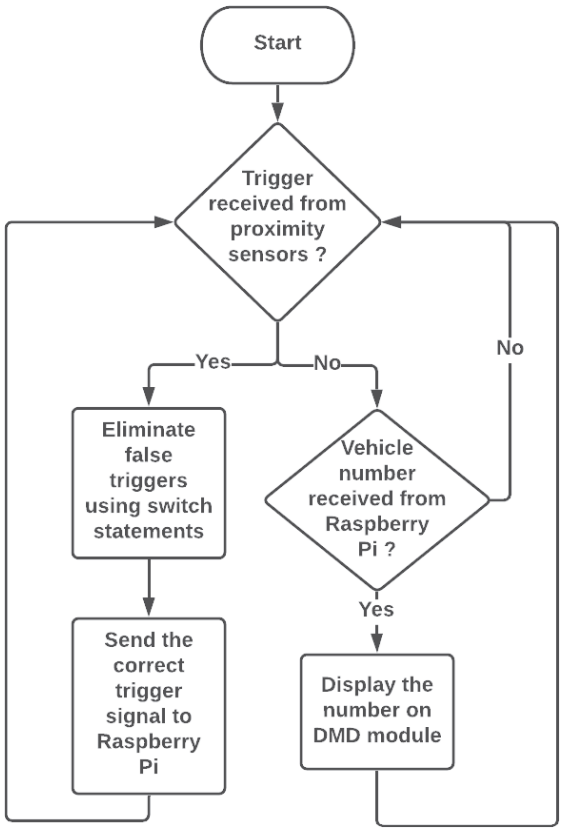


Fig.4 Arduino Process Flow Diagram

IV. EXPERIMENTAL RESULTS

We performed a simulation of Arduino Uno and DMD display using Proteus PCB Design and Circuit Simulation software [Fig.5]. For interfacing the DMD display with Arduino, we have used an array of 10 MAX7219 ICs. To simulate the serial communication between the microcontroller and Raspberry Pi, we have used a virtual terminal through which the vehicle number is being provided for the DMD display panel.

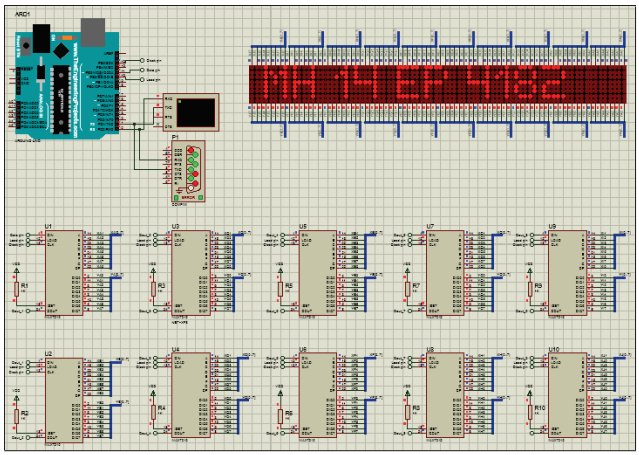


Fig.5 Proteus Simulation of Arduino & DMD module

The vehicle detection algorithm based on MobilenetSSD when trained on the PASCAL VOC 2007+2012 dataset has an overall mean average precision metric of 74.3. Focusing on individual object classes, the mAP for the object class 'car' is 76.1, while that of 'motorbike' is 82.3. Tesseract based Optical Character Recognition has a character accuracy of 70% for gray scale images [13].

To evaluate the OCR performance, we prepared a test dataset of Indian vehicle number plates comprising of 1200 images, out of which 413 were recognized correctly, thus, indicating an impractical accuracy rate of 34.41%. While the algorithm proposed by Rayson Laroca et al. has an accuracy of 78.33% [11] on the UFPR-ALPR dataset, whereas, the accuracy on testing with our Indian number plate dataset is 65.58%.

V. CONCLUSIONS

The Wrong Side Vehicle Detection system has been successfully implemented using low cost hardware comprising IR proximity sensors, processing unit ,i.e., Arduino, Raspberry Pi, Camera and DMD display. The prototype model exhibits satisfactory results which are reproducible in the Indian road conditions. Moreover, the model is fixed on a frame of manageable dimensions weighing 15 kg, allowing portability and can be easily setup by an individual.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:59:34 UTC from IEEE Xplore. Restrictions apply.



Fig.6 Testing results. The first stage shows vehicle being detected upon receiving trigger, the number plate is being detected in the second stage and cropped for OCR

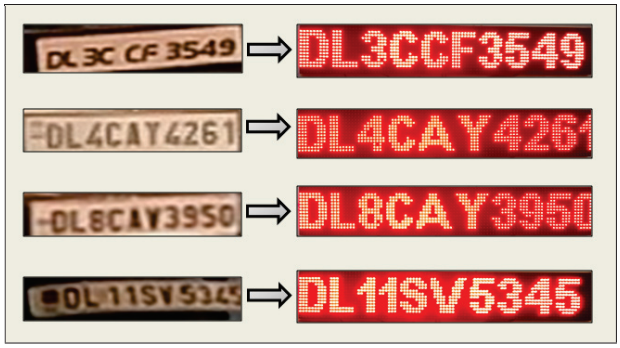


Fig.7 Number displayed on DMD after OCR

VI. Discussions

The vehicle detection algorithm based on MobilenetSSD object detection has an overall mean average precision metric of 74.3. Focusing on individual object classes, the mAP for the object class 'car' is 76.1, while that of 'motorbike' is 82.3. Parameters such as speed of the vehicle can hamper the detection efficiency which might lead to erroneous results. The UPFR-ALPR dataset used in [11] for training the YOLO based OCR model comprises of 4500 images of vehicles wherein, the registration numbers are made of 3 letters and 4 digits. On the other hand, our dataset consists of 1200 images of Indian vehicle number plates that have a length ranging from 9 to 11 characters, a typical number would have 5 letters and 5 digits. By curating a comprehensive data set of vehicle number plates under varying conditions, the OCR algorithm can be improved further.

The commercially available IR proximity sensor used for triggering the vehicle detection algorithm has a maximum detection range of 85 cm. If the vehicle moving in the wrong direction approaches the sensor at a distance beyond the maximum range then it may go undetected. For large scale implementation of the system, custom built proximity sensors can be used with an enhanced range of 1.5 to 2 meters. However this may result in increased costs and power consumption.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:59:34 UTC from IEEE Xplore. Restrictions apply.

TABLE III
EXPERIMENTAL RESULTS

Method	mAP
MobileNet based Vehicle Detection	74.3
YOLO based OCR on UFPR-ALPR[11]	78.33
YOLO based OCR on our dataset	65.58
Tesseract based OCR on our dataset	34

The scope of wrong side vehicle detection system can be expanded by implementing vehicle to infrastructure (V2I) systems. To alert other road users, information can be directly relayed through wireless communication, bypassing the LED display. Moreover, the system can be employed for detecting traffic signal and speed limit violators by using vehicle tracking. The object detection system can be employed to gauge traffic volume, thus enabling smart traffic control. The information extracted from the vehicle license plate can be sent to traffic authorities for penalising the offenders and further analysis.

REFERENCES

[1] R. V. Ponnaluri, "Modeling wrong-way crashes and fatalities on arterials and freeways," *IATSS Research*, vol. 42, no. 1, pp. 8–17, 2018.

[2] G. Van Rossum and F. L. Drake, *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009.

[3] S. Usmankhujaev, S. Baydadaev, and K. Woo, "Real-time, deep learning based wrong direction detection," *Applied Sciences*, vol. 10, p. 2453, 04 2020.

[4] P. Kulkarni, A. Khatri, P. Banga, and K. Shah, "Automatic number plate recognition (anpr) system for indian conditions," in *2009 19th International Conference Radioelektronika*, 2009, pp. 111–114.

[5] A. Beibut, K. Magzhan, and K. Chingiz, "Effective algorithms and methods for automatic number plate recognition," in *2014 IEEE 8th International Conference on Application of Information and Communication Technologies (AICT)*, 2014, pp. 1–4.

[6] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," 2017.

[7] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," *Lecture Notes in Computer Science*, p. 21–37, 2016. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-46448-0_2

[8] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788.

[9] Z. Rahman, A. M. Ami, and M. A. Ullah, "A real-time wrong-way vehicle detection based on yolo and centroid tracking," in *2020 IEEE Region 10 Symposium (TENSYMP)*, 2020, pp. 916–920.

[10] R. Smith, "An overview of the tesseract ocr engine," in *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, vol. 2, 2007, pp. 629–633.

[11] R. Laroca, E. Severo, L. A. Zanlorensi, L. S. Oliveira, G. R. Gonçalves, W. R. Schwartz, and D. Menotti, "A robust real-time automatic license plate recognition based on the YOLO detector," in *International Joint Conference on Neural Networks (IJCNN)*, July 2018, pp. 1–10.

[12] S. Montazzolli and C. Jung, "Real-time brazilian license plate detection and recognition using deep convolutional neural networks," 10 2017.

[13] C. Patel, A. Patel, and D. Patel, "Optical character recognition by open source ocr tool tesseract: A case study," in *Int. J. Comput. Appl.*, vol. 55, no. 10, 2012, pp. 50–56.

[14] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

[15] M. Piyaneeranart and M. Ketcham, "Detection of moving vehicle spacing," in *2018 18th International Symposium on Communications and Information Technologies (ISCIT)*, 2018, pp. 156–159.

[16] G. Monteiro, M. Ribeiro, J. Sena Marcos, and J. Batista, "Wrongway drivers detection based on optical flow," vol. 5, 01 2007, pp. V – 141.

[17] S. Ozkul and P.-S. Lin, "Evaluation of red rrfb implementation at freeway off-ramps and its effectiveness on alleviating wrong-way driving," *Transportation Research Procedia*, vol. 22, pp. 570–579, 12 2017.

[18] F. Baratian-Ghorghi and H. Zhou, "Traffic control devices for deterring wrong-way driving: Historical evolution and current practice," *Journal of Traffic and Transportation Engineering (English Edition)*, vol. 4, no. 3, pp. 280–289, 2017.

[19] A. Mangal, A. Goyal, and A. R. Kulkarni, "Design and development of iot based prototype police barricade," in *2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*, 2020, pp. 1–6.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on September 01,2021 at 06:59:34 UTC from IEEE Xplore. Restrictions apply.